Anton Bovier

Gaussian processes on trees: From spin glasses to branching Brownian motion

Lecture Notes, Bonn, WS 2014/15

March 11, 2015
Preface

These are notes written for a graduate course at Bonn University I gave in 2014/15. It extends a shorter version of lecture notes for a series of lectures given at the 7th Prague Summer School on Mathematical Statistical Mechanics in August 2013.

BBM was studied over the last 50 years as a subject of interest in its own right, with seminal contributions by McKean, Bramson, Lalley, Sellke, Chauvin, Rouault, and others. Recently, the field has experienced a revival with many remarkable contributions and repercussions into other areas. The construction of the extremal process in our work with Arguin and Kistler [6] as well as, in parallel, in that of Aïdékon, Beresticky, Brunet, and Shi [2], is one highlight. Many results on BBM were extended to branching random walk, see e.g. Aïdékon [1]. Other developments concern the connection to extremes of the free Gaussian random field in $d = 2$ by Bramson and Zeitouni [25], Ding [32], Bramson, Ding, and Zeitouni [24], and Biskup and Louidor [13, 14]. My personal motivation to study this object came, however, from spin glass theory. About ten years ago, Irina Kurkova and I wrote two papers [19, 20] on Derrida’s Generalised Random Energy models (GREM). It turns out that these models are closely related to BBM, and with BBM corresponding to a critical case, that we could not at the time fully analyse.

These lecture notes are mainly motivated by the work we did with Nicola Kistler, Louis-Pierre Arguin, and Lisa Hartung. The aim of the course is to give a comprehensive picture of what goes into the analysis of the extremal properties of branching Brownian motions, seen as Gaussian processes indexed by Galton-Watson trees. Chapters 1-3 provide some standard background material on extreme value theory, point processes, and Gaussian processes. Chapter 4 gives a brief glimpse on spin glasses, in particular the REM and GREM models of Derrida which provides some important motivation. Chapter 5 introduces branching Brownian motion and its relation to the Fischer-Kolmogorov-Petrovsky-Piscounov equations and states some by now classical results of Bramson and of Lalley and Sellke. Chapter 6 gives a condensed but detailed review of the analysis of the F-KPP equations that was given in Bramson’s monograph [22]. The remainder of the notes is devoted to more recent work. In Chapter 7 I review the derivation and description of the extremal process of
BBM contained mostly the paper [6] with Louis-Pierre Arguin and Nicola Kistler. Chapter 8 describes recent work with Lisa Hartung [16] on an extension of that work. Chapter 9 reviews two papers [17, 18] with Lisa Hartung on variable speed branching Brownian motion. Discussion of recent work on related problems, such as branching random walks, the Gaussian free field, and others, may possibly be added in later versions.

The recent activities in and around BBM have triggered a number of lecture notes and reviews, that hopefully are complementary to this one. I mention the review by Gouéré [40] that presents and compares the approaches of Arguin et al and Aïdékon et al.. Ofer Zeitouni has lecture notes on his homepage that deal with branching random walk and the Gaussian free field [75], Nicola Kistler just wrote a survey linking REM and GREM to other correlated Gaussian processes [53].

I am deeply grateful to my collaborators on the matters of these lectures, Louis-Pierre Arguin, Nicola Kistler, Irina Kurkova, and Lisa Hartung. They did the bulk of the work, and without them none this would have been achieved.

I also thank Jiří Černý and Lisa Hartung for pointing out various mistakes in previous versions of these notes. I thank Marek Biskup, Jiří Černý, and Roman Kotecký for organising the excellent school in Prague that ultimately triggered the idea to produce such notes.

Bonn, February 2015,  
Anton Bovier
Contents

1 Extreme value theory for iid sequences .............................................. 1
  1.1 Basic issues ................................................................. 1
  1.2 Extremal distributions ...................................................... 2
    1.2.1 Example: The Gaussian distribution .............................. 3
    1.2.2 Some technical preparation ........................................ 4
    1.2.3 max-stable distributions ........................................... 8
    1.2.4 The extremal types theorem ....................................... 10
  1.3 Level-crossings and the distribution of the $k$-th maxima .......... 11

2 Extremal processes ................................................................. 15
  2.1 Point processes .............................................................. 15
  2.2 Laplace functionals .......................................................... 18
  2.3 Poisson point processes ..................................................... 19
  2.4 Convergence of point processes .......................................... 20
    2.4.1 Vague convergence .................................................. 21
    2.4.2 Weak convergence .................................................... 23
  2.5 Point processes of extremes .............................................. 28
    2.5.1 The point process of exceedances ................................. 28
    2.5.2 The point process of extreme values ............................. 29
    2.5.3 Complete Poisson convergence ................................... 30

3 Normal sequences ................................................................. 33
  3.1 Normal comparison .......................................................... 34
  3.2 Applications to extremes .................................................. 39

4 Spin glasses ....................................................................... 43
  4.1 Setting and examples ....................................................... 43
  4.2 The REM ................................................................. 45
    4.2.1 Rough estimates, the second moment method .................. 45
    4.2.2 Law of the maximum and the extremal process ............... 47
  4.3 The GREM, two levels ..................................................... 47


<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.1</td>
<td>The finer computations</td>
<td>49</td>
</tr>
<tr>
<td>4.4</td>
<td>The general model and relation to branching Brownian motion</td>
<td>51</td>
</tr>
<tr>
<td>4.5</td>
<td>The Galton-Watson process</td>
<td>52</td>
</tr>
<tr>
<td>4.6</td>
<td>The REM on the Galton-Watson tree</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>Branching Brownian motion</td>
<td>57</td>
</tr>
<tr>
<td>5.1</td>
<td>Definition and basics</td>
<td>57</td>
</tr>
<tr>
<td>5.2</td>
<td>Rough heuristics</td>
<td>58</td>
</tr>
<tr>
<td>5.3</td>
<td>Recursion relations</td>
<td>60</td>
</tr>
<tr>
<td>5.4</td>
<td>The F-KPP equation</td>
<td>61</td>
</tr>
<tr>
<td>5.5</td>
<td>The travelling wave</td>
<td>63</td>
</tr>
<tr>
<td>5.6</td>
<td>The derivative martingale</td>
<td>66</td>
</tr>
<tr>
<td>6</td>
<td>Bramson's analysis of the F-KPP equation</td>
<td>71</td>
</tr>
<tr>
<td>6.1</td>
<td>Feynman-Kac representation</td>
<td>71</td>
</tr>
<tr>
<td>6.2</td>
<td>The maximum principle and its applications</td>
<td>75</td>
</tr>
<tr>
<td>6.3</td>
<td>Estimates on solutions of the linear F-KPP equation</td>
<td>89</td>
</tr>
<tr>
<td>6.4</td>
<td>Brownian bridges</td>
<td>92</td>
</tr>
<tr>
<td>6.5</td>
<td>Hitting probabilities of curves</td>
<td>95</td>
</tr>
<tr>
<td>6.6</td>
<td>Asymptotics of solutions of the F-KPP equation</td>
<td>97</td>
</tr>
<tr>
<td>6.7</td>
<td>Convergence results</td>
<td>104</td>
</tr>
<tr>
<td>7</td>
<td>The extremal process of BBM</td>
<td>115</td>
</tr>
<tr>
<td>7.1</td>
<td>Limit theorems for solutions</td>
<td>115</td>
</tr>
<tr>
<td>7.2</td>
<td>Existence of a limiting process</td>
<td>120</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Convergence of the Laplace functional</td>
<td>120</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Flash back to the derivative martingale</td>
<td>122</td>
</tr>
<tr>
<td>7.2.3</td>
<td>A representation for the Laplace functional</td>
<td>123</td>
</tr>
<tr>
<td>7.3</td>
<td>Interpretation as cluster point process</td>
<td>124</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Interpretation via an auxiliary process</td>
<td>124</td>
</tr>
<tr>
<td>7.3.2</td>
<td>The Poisson process of cluster extremes</td>
<td>126</td>
</tr>
<tr>
<td>7.3.3</td>
<td>More on the law of the clusters</td>
<td>128</td>
</tr>
<tr>
<td>7.3.4</td>
<td>The extremal process seen from the cluster extremes</td>
<td>130</td>
</tr>
<tr>
<td>8</td>
<td>Full extremal process</td>
<td>135</td>
</tr>
<tr>
<td>8.1</td>
<td>Introduction</td>
<td>135</td>
</tr>
<tr>
<td>8.1.1</td>
<td>The embedding</td>
<td>135</td>
</tr>
<tr>
<td>8.1.2</td>
<td>The extended convergence result</td>
<td>136</td>
</tr>
<tr>
<td>8.2</td>
<td>Properties of the embedding</td>
<td>137</td>
</tr>
<tr>
<td>8.3</td>
<td>The q-thinning</td>
<td>143</td>
</tr>
<tr>
<td>9</td>
<td>Variable speed BBM</td>
<td>149</td>
</tr>
<tr>
<td>9.1</td>
<td>The construction</td>
<td>149</td>
</tr>
<tr>
<td>9.2</td>
<td>Two-speed BBM</td>
<td>150</td>
</tr>
<tr>
<td>9.2.1</td>
<td>Position of extremal particles at time $bt$</td>
<td>152</td>
</tr>
<tr>
<td>9.2.2</td>
<td>Convergence of the McKean martingale</td>
<td>154</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.2.3</td>
<td>Asymptotic behaviour of BBM</td>
<td>158</td>
</tr>
<tr>
<td>9.2.4</td>
<td>Convergence of the maximum</td>
<td>159</td>
</tr>
<tr>
<td>9.2.5</td>
<td>Existence of the limiting process</td>
<td>165</td>
</tr>
<tr>
<td>9.2.6</td>
<td>The auxiliary process</td>
<td>165</td>
</tr>
<tr>
<td>9.2.7</td>
<td>The case $\sigma_1 &gt; \sigma_2$</td>
<td>169</td>
</tr>
<tr>
<td>9.3</td>
<td>Universality below the straight line</td>
<td>172</td>
</tr>
<tr>
<td>9.3.1</td>
<td>Outline of the proof</td>
<td>173</td>
</tr>
<tr>
<td>9.3.2</td>
<td>Localization of paths</td>
<td>174</td>
</tr>
<tr>
<td>9.3.3</td>
<td>Proof of Proposition 9.29</td>
<td>179</td>
</tr>
<tr>
<td>9.3.4</td>
<td>Approximating two-speed BBM.</td>
<td>180</td>
</tr>
<tr>
<td>9.3.5</td>
<td>Gaussian comparison</td>
<td>181</td>
</tr>
</tbody>
</table>

References ...................................................................... 187

Index ...................................................................................... 191
Chapter 1
Extreme value theory for iid sequences

The goal of these lectures is the analysis of extremes in a large class of Gaussian processes on trees. To get this into perspective, we need to start with the simplest and standard situation, the theory of extremes of sequences of independent random variables. This is the subject of this opening chapter.

Records and extremes are not only fascinating us in all areas of live, they are also of tremendous importance. We are constantly interested in knowing how big, how, small, how rainy, how hot, etc. things may possibly be. To answer such questions, an entire branch of statistics, called extreme value statistics, was developed. There are a large number of textbooks on this subject, my personal favourites being those by Leadbetter, Lindgren, and Roozen [57] and by Resnick [65].

1.1 Basic issues

As usual in statistics, one starts with a set of observations, or data, that correspond to partial observations of some sequence of events. Let us assume that these events are related to the values of some random variables, $X_i, i \in \mathbb{Z}$, taking values in the real numbers. Assuming that $\{X_i\}_{i \in \mathbb{Z}}$ is a stochastic process (with discrete time) defined on some probability space $(\Omega, \mathcal{F}, P)$, our first question will be about the distribution of its maximum: Given $n \in \mathbb{N}$, define the maximum up to time $n$,

$$M_n \equiv \max_{i=1}^n X_i. \quad (1.1.1)$$

We then ask for the distribution of this new random variable, i.e. we ask what is $P(M_n \leq x)$? As often, one is interested in this question particularly when $n$ is large, i.e. we are interested in the asymptotics as $n \uparrow \infty$.

Certainly $M_n$ may tend to infinity, and their distribution may have no reasonable limit. The natural first question about $M_n$ are thus: first, can we rescale $M_n$ in some way such that the rescaled variable converges to a random variable, and second, is
1 Extreme value theory for iid sequences

there a universal class of distributions that arises as the distribution of the limits? To answer these questions will be our first target.

A second major issue will be to go beyond just the maximum value. Coming back to the marks of flood levels under the bridge, we do not just see one, but a whole bunch of marks. Can we say something about their joint distribution? In other words, what is the law of the maximum, the second largest, third largest, etc.? Is there, possibly again a universal law of how this process of extremal marks looks like? This will be the second target, and we will see that there is again an answer to the affirmative.

1.2 Extremal distributions

We consider a family of real valued, independent identically distributed random variables \(X_i, i \in \mathbb{N}\), with common distribution function \(F(x) \equiv \mathbb{P}(X_i \leq x)\). (1.2.1)

Recall that by convention, \(F(x)\) is a non-decreasing, right-continuous function \(F: \mathbb{R} \to [0, 1]\). Note that the distribution function of \(M_n\) is simply

\[
\mathbb{P}(M_n \leq x) = \mathbb{P}(\forall_{i=1}^n X_i \leq x) = \prod_{i=1}^n \mathbb{P}(X_i \leq x) = (F(x))^n. (1.2.2)
\]

As \(n\) tends to infinity, this will converge to a trivial limit

\[
\lim_{n \to \infty} (F(x))^n = \begin{cases} 0, & \text{if } F(x) < 1, \\ 1, & \text{if } F(x) = 1. \end{cases} (1.2.3)
\]

which simply says that any value that the variables \(X_i\) can exceed with positive probability will eventually exceeded after sufficiently many independent trials.

As we have already indicated above, to get something more interesting, we must rescale. It is natural to try something similar to what is done in the central limit theorem: first subtract an \(n\)-dependent constant, then rescale by an \(n\)-dependent factor. Thus the first question is whether one can find two sequences, \(b_n\), and \(a_n\), and a non-trivial distribution function, \(G(x)\), such that

\[
\lim_{n \to \infty} \mathbb{P}(a_n(M_n - b_n)) = G(x). (1.2.4)
\]
1.2 Extremal distributions

1.2.1 Example: The Gaussian distribution.

In probability theory, it is always natural to start playing with the example of a Gaussian distribution. So we now assume that our \( X_i \) are Gaussian, i.e. that \( F(x) = \Phi(x) \), where

\[
\phi(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy.
\]

(1.2.5)

We want to compute

\[
\mathbb{P}(a_n(M_n - b_n) \leq x) = \mathbb{P}(M_n \leq a_n^{-1} x + b_n) = (\Phi(a_n^{-1} x + b_n))^n.
\]

(1.2.6)

Setting \( x_n \equiv a_n^{-1} x + b_n \), this can be written as

\[
(1 - (1 - \Phi(x_n))^n)
\]

(1.2.7)

For this to converge, we must choose \( x_n \) such that

\[
(1 - \Phi(x_n)) = n^{-1} g(x) + o(1/n),
\]

in which case

\[
\lim_{n \to \infty} (1 - (1 - \Phi(x_n))^n) = e^{-g(x)}.
\]

(1.2.8)

(1.2.9)

Thus our task is to find \( x_n \) such that

\[
\frac{1}{\sqrt{2\pi}} \int_{x_n}^{\infty} e^{-y^2/2} dy. = n^{-1} g(x).
\]

(1.2.10)

At this point it will be very convenient to use an approximation for the function \( 1 - \Phi(u) \) when \( u \) is large, namely

\[
\frac{1}{u\sqrt{2\pi}} e^{-u^2/2} (1 - (2u^{-2})) \leq 1 - \Phi(u) \leq \frac{1}{u\sqrt{2\pi}} e^{-u^2/2}. \]

(1.2.11)

Note the these bounds area going to be used over and over, and are surely worth memorising. Using this bound, our problem simplifies to solving

\[
\frac{1}{x_n\sqrt{2\pi}} e^{-x_n^2/2} = n^{-1} g(x),
\]

(1.2.12)

that is

\[
n^{-1} g(x) = \frac{e^{-\frac{1}{2}(a_n^{-1} x + b_n)^2}}{\sqrt{2\pi} a_n^{-1} x + b_n} = \frac{e^{-b_n^2/2 - a_n^{-2} t^2/2 - a_n^{-1} b_n x}}{\sqrt{2\pi}(a_n^{-1} x + b_n)}. \]

(1.2.13)

Setting \( x = 0 \), we find

\[
\frac{e^{-b_n^2/2}}{\sqrt{2\pi} b_n} = n^{-1} g(0).
\]

(1.2.14)
Let us make the ansatz $b_n = \sqrt{2\ln n} + c_n$. Then we get for $c_n$
\[ e^{-\sqrt{2\ln n} - c_n^2/2} = \frac{1}{\sqrt{2\pi}}(\sqrt{2\ln n} + c_n). \] (1.2.15)

It is convenient to choose $g(0) = 1$. Then, the leading terms for $c_n$ are given by
\[ c_n = -\frac{\ln \ln n + \ln(4\pi)}{2\sqrt{2\ln n}}. \] (1.2.16)

The higher order corrections to $c_n$ can be ignored, as they do not affect the validity of (1.2.8). Finally, inspecting (1.2.13), we see that we can choose $a_n = \sqrt{2\ln n}$. Putting all things together we arrive at the following assertion.

**Lemma 1.1.** Let $X_i, i \in \mathbb{N}$ be iid normal random variables. Let
\[ b_n \equiv \sqrt{2\ln n} - \frac{\ln \ln n + \ln(4\pi)}{2\sqrt{2\ln n}}, \] (1.2.17)
and
\[ a_n = \sqrt{2\ln n}. \] (1.2.18)

Then, for any $x \in \mathbb{R}$,
\[ \lim_{n \to \infty} \mathbb{P}(a_n(M_n - b_n) \leq x) = e^{-e^{-x}}. \] (1.2.19)

**Remark 1.2.** It will be sometimes convenient to express (1.2.19) in a slightly different, equivalent form. With the same constants, $a_n, b_n$, define the function
\[ u_n(x) \equiv b_n + x/a_n. \] (1.2.20)

Then
\[ \lim_{n \to \infty} \mathbb{P}(M_n \leq u_n(x)) = e^{-e^{-x}}. \] (1.2.21)

This is our first result on the convergence of extremes, and the function $e^{-e^{-x}}$, called the **Gumbel distribution**, is the first extremal distribution that we encounter.

The next question to ask is how “typical” the result for the Gaussian distribution is. From the computation we see readily that we made no use of the Gaussian hypothesis to get the general form $\exp(-g(x))$ for any possible limit distribution. The fact that $g(x) = \exp(-x)$, however, depended on the particular form of $\Phi$. We will see next that, remarkably, only two other types of functions can occur.

### 1.2.2 Some technical preparation.

Our goal will be to be as general as possible with regard to the allowed distributions $F$. Of course we must anticipate that in some cases, no limiting distributions can be
constructed (e.g. think of the case of a distribution with support on the two points 0 and 1!). Nonetheless, we are not willing to limit ourselves to random variables with continuous distribution functions, and this will introduce a little bit of complication, that, however, can be seen as a useful exercise.

Before we continue, let us explain where we are heading. In the Gaussian case we have seen already that we could make certain choices at various places. In general, we can certainly multiply the constants \( a_n \) by a finite number and add a finite number to the choice of \( b_n \). This will clearly result in a different form of the extremal distribution, which, however, we think as morally equivalent. Thus, when classifying extremal distributions, we will think of two distributions, \( G, F \), as equivalent if

\[
F(ax + b) = G(x). \tag{1.2.22}
\]

The distributions we are looking for arise as limits of the form

\[
F^n(a_n x + b_n) \to G(x). \tag{1.2.23}
\]

We will want to use that such limits have particular properties, namely that for some choices of \( \alpha_n, \beta_n \),

\[
G^n(\alpha_n x + \beta_n) = G(x). \tag{1.2.24}
\]

This property is called max-stability. Our program is then reduced to classify all max-stable distributions modulo the equivalence (1.2.22), and to determine their domains of attraction. Note the similarity of the characterisation of the Gaussian distribution as a stable distribution under addition of random variables.

Recall the notion of weak convergence of distribution functions:

**Definition 1.3.** A sequence, \( F_n \), of probability distribution functions is said converge weakly to a probability distribution function \( F \),

\[
F_n \overset{w}{\to} F, \tag{1.2.25}
\]

iff and only if

\[
F_n(x) \to F(x), \tag{1.2.26}
\]

for all points \( x \) where \( F \) is continuous.

The next thing we want to do is to define the notion of the (left-continuous) inverse of a non-decreasing, right-continuous function (that may have jumps and flat pieces).

**Definition 1.4.** Let \( \psi : \mathbb{R} \to \mathbb{R} \) be a monotone increasing, right-continuous function. Then the inverse function \( \psi^{-1} \) is defined as

\[
\psi^{-1}(y) \equiv \inf \{ x : \psi(x) \geq y \}. \tag{1.2.27}
\]

We need the following properties of \( \psi^{-1} \).

**Lemma 1.5.** Let \( \psi \) be as in the definition, and \( a > c \) and \( b \) real constants. Let \( H(x) \equiv \psi(ax + b) - c \). Then
Extreme value theory for iid sequences

\( G(i) \)

(i) \( \psi^{-1} \) is left-continuous.
(ii) \( \psi(\psi^{-1}(x)) \geq x \).
(iii) If \( \psi^{-1} \) is continuous at \( \psi(x) \in \mathbb{R} \), then \( \psi^{-1}(\psi(x)) = x \).
(iv) \( H^{-1}(y) = a^{-1}(\psi^{-1}(y+c) - b) \)
(v) If \( G \) is a non-degenerate distribution function, then there exist \( y_1 < y_2 \), such that \( G^{-1}(y_1) < G^{-1}(y_2) \).

Proof. (i) First note that \( \psi^{-1} \) is increasing. Let \( y_n \uparrow y \). Assume that \( \lim_n \psi^{-1}(y_n) < \psi^{-1}(y) \). This means that for all \( y_n \), \( \inf\{x : \psi(x) \geq y_n\} < \inf\{x : \psi(x) \geq y\} \). This means that there is a number, \( x_0 < \psi^{-1}(y) \), such that, for all \( n \), \( \psi(x_0) \leq y_n \), but \( \psi(x_0) > y \). But this means that \( \lim_n y_n \geq y \), which is in contradiction to the hypothesis. Thus \( \psi^{-1} \) is left-continuous.
(ii) is immediate from the definition.
(iii) \( \psi^{-1}(\psi(x)) = \inf\{x' : \psi(x') \geq \psi(x)\} \), thus obviously \( \psi^{-1}(\psi(x)) \leq x \). On the other hand, for any \( \varepsilon > 0 \), \( \psi^{-1}(\psi(x) + \varepsilon) = \inf\{x' : \psi(x') \geq \psi(x) + \varepsilon\} \). But \( \psi(x') \) can only be strictly greater than \( \psi(x) \) if \( x' > x \), so for any \( y' > \psi(x) \), \( \psi^{-1}(y') \geq x \). Thus, if \( \psi^{-1} \) is continuous at \( \psi(x) \), this implies that \( \psi^{-1}(\psi(x)) = x \).
(iv) The verification of the formula for the inverse of \( H \) is elementary and left as an exercise.
(v) If \( G \) is not degenerate, then there exist \( x_1 < x_2 \) such that \( 0 < G(x_1) \equiv y_1 < G(x_2) \equiv y_2 \leq 1 \). But then \( G^{-1}(y_1) \leq x_1 \), and \( G^{-1}(y_2) = \inf\{x : G(x) \geq G(x_2)\} \). If the latter equals \( x_1 \), then for all \( x' \geq x_1 \), \( G(x') \geq G(x_2) \), and since \( G \) is right-continuous, \( G(x_1) = G(x_2) \), which is a contradiction. □

The following corollary is important.

Corollary 1.6. If \( G \) is a non-degenerate distribution function, and there are constants \( a > 0 \), \( \alpha > 0 \), and \( b, \beta \in \mathbb{R} \), such that, for all \( x \in \mathbb{R} \),

\[
G(ax + b) = G(ax + \beta).
\] (1.2.28)

then \( a = \alpha \) and \( b = \beta \).

Proof. Set \( H(x) \equiv G(ax + b) \). Then, by (i) of the preceding lemma,

\[
H^{-1}(y) = a^{-1}(G^{-1}(y) - b),
\] (1.2.29)

but by (1.2.28) also

\[
H^{-1}(y) = \alpha^{-1}(G^{-1}(y) - \beta).
\] (1.2.30)

On the other hand, by (v) of the same lemma, there are at least two values of \( y \) such that \( G^{-1}(y) \) are different, i.e. there are \( x_1 < x_2 \) such that

\[
a^{-1}(x_1 - b) = \alpha^{-1}(x_i - \beta),
\] (1.2.31)

which obviously implies the assertion of the corollary. □

Remark 1.7. Note that the assumption that \( G \) is non-degenerate is necessary. If, e.g., \( G(x) \) has a single jump from 0 to 1 at a point \( a \), then it holds that \( G(5x - 4a) = G(x) \)!
1.2 Extremal distributions

The next theorem is known as Khintchine’s theorem:

**Theorem 1.8.** Let $F_n, n \in \mathbb{N}$, be distribution functions, and let $G$ be a non-degenerate distribution function. Let $a_n > 0$, and $b_n \in \mathbb{R}$ be sequences such that

$$F_n(a_nx + b_n) \xrightarrow{w} G(x). \quad (1.2.32)$$

Then there are constants, $\alpha_n > 0$ and $\beta_n \in \mathbb{R}$, and a non-degenerate distribution function, $G_*$, such that

$$F_n(\alpha_nx + \beta_n) \xrightarrow{w} G_*(x), \quad (1.2.33)$$

if and only if

$$a_n^{-1} \alpha_n \rightarrow a, \quad (\beta_n - b_n)/a_n \rightarrow b, \quad (1.2.34)$$

and

$$G_*(x) = G(ax + b). \quad (1.2.35)$$

**Remark 1.9.** This theorem makes the comment made above precise, saying that different choices of the scaling sequences $a_n, b_n$ can lead only to distributions that are related by a transformation $(1.2.35)$.

**Proof.** By changing $F_n$, we can assume for simplicity that $a_n = 1$, $b_n = 0$. Let us first show that if $\alpha_n \rightarrow a, \beta_n \rightarrow b$, then $F_n(\alpha_nx + \beta_n) \rightarrow G_*(x)$.

Let $ax + b$ be a point of continuity of $G$. Write

$$F_n(\alpha_nx + \beta_n) = F_n(\alpha_nx + \beta_n) - F_n(ax + b) + F_n(ax + b). \quad (1.2.36)$$

By assumption, the last term converges to $G(ax + b)$. Without loss of generality we may assume that $\alpha_nx + \beta_n$ is monotone increasing. We want to show that

$$F_n(\alpha_nx + \beta_n) - F_n(ax + b) \uparrow 0. \quad (1.2.37)$$

For, otherwise, there would be a constant, $\delta > 0$, such that along a subsequence $n_k$, $\lim_k F_n(\alpha_{n_k}x + \beta_{n_k}) - F_n(ax + b) < -\delta$. But since $\alpha_{n_k}x + \beta_{n_k} \uparrow ax + b$, this implies that for any $y < ax + b$, $\lim_k F_n(y) - F_n(ax + b) < -\delta$. Now, if $G$ is continuous at $y$, this implies that $G(y) - G(ax + b) < -\delta$. But this implies that either $F$ is discontinuous at $ax + b$, or there exists a neighbourhood of $ax + b$ such that $G(x)$ has no point of continuity within this neighbourhood. But this is impossible since a probability distribution function can only have countably many points of discontinuity. Thus $(1.2.37)$ must hold, and hence

$$F_n(\alpha_nx + \beta_n) \xrightarrow{w} G(ax + b) \quad (1.2.38)$$

which proves $(1.2.33)$ and $(1.2.35)$.

Next we want to prove the converse, i.e. we want to show that $(1.2.33)$ implies $(1.2.34)$. Note first that $(1.2.33)$ implies that the sequence $\alpha_nx + \beta_n$ is bounded, since otherwise there would be subsequences converging to plus or minus infinity, along those $F_n(\alpha_nx + \beta_n)$ would converge to 0 or 1, contradicting the assumption. This implies that the sequence has converging subsequences, $\alpha_{n_k}, \beta_{n_k}$, along which
Then the preceding results shows that $a_{nk} \rightarrow a', b_{nk} \rightarrow b'$, and $G_{a'}(x) = G(a'x + b')$. Now, if the sequence does not converge, there must be another convergent subsequence $a_{nk}' \rightarrow a''$, $b_{nk}' \rightarrow b''$. But then

$$G_{a'}(x) = \lim_{k} F_{a''_{nk}}(\alpha_{nk}' x + \beta_{nk}') \rightarrow G(a''x + b').$$

Thus $G(a'x + b') = G(a''x + b'')$. and so, since $G$ is non-degenerate, Corollary 1.6 implies that $a' = a''$ and $b' = b''$, contradicting the assumption that the sequences do not converge. This proves the theorem. $\square$

1.2.3 Max-stable distributions.

We are now prepared to continue our search for extremal distributions. Let us formally define the notion of max-stable distributions.

**Definition 1.10.** A non-degenerate probability distribution function, $G$, is called **max-stable**, if for all $n \in \mathbb{N}$, there exists $a_n > 0, b_n \in \mathbb{R}$, such that, for all $x \in \mathbb{R}$,

$$G^n(a_n^{-1}x + b_n) = G(x).$$

(1.2.41)

The next proposition gives some important equivalent formulations of max-stability and justifies the term.

**Proposition 1.11.** (i) A probability distribution, $G$, is max-stable, if and only if there exists probability distributions $F_n$ and constants $a_n > 0, b_n \in \mathbb{R}$, such that, for all $k \in \mathbb{N}$,

$$F_{a_{nk}^{-1}x + b_{nk}} \overset{w}{\rightarrow} G^{1/k}(x).$$

(1.2.42)

(ii) $G$ is max-stable if and only if there exists a probability distribution function, $F$, and constants $a_n > 0, b_n \in \mathbb{R}$, such that

$$F^n(a_n^{-1}x + b_n) \overset{w}{\rightarrow} G(x).$$

(1.2.43)

**Proof.** We first prove (i). If (1.2.42) holds, then by Khintchine’s theorem, there exist constants, $\alpha_k, \beta_k$, such that

$$G^{1/k}(x) = G(\alpha_k x + \beta_k),$$

(1.2.44)

for all $k \in \mathbb{N}$, and thus $G$ is max-stable. Conversely, if $G$ is max-stable, set $F_n = G^n$, and let $a_n, b_n$ the constants that provide for (1.2.41). Then

$$F_n(a_{nk}^{-1}x + b_{nk}) = \left[G^n(a_{nk}^{-1}x + b_{nk})\right]^{1/k} = G^{1/k},$$

(1.2.45)
1.2 Extremal distributions

which proves the existence of the sequence $F_n$ and of the respective constants.

Now let us prove (ii). Assume first that $G$ is max-stable. Then choose $F = G$. Then the fact that $\lim_n F^n(a_n^{-1}x + b_n) = G(x)$ follows if the constants from the definition of max-stability are used trivially.

Next assume that (1.2.43) holds. Then, for any $k \in N$,

\[
F^{nk}(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G(x),
\]

and so

\[
F^n(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G^{1/k}(x),
\]

so $G$ is max-stable by (i)! \qed

There is a slight extension to this result.

**Corollary 1.12.** If $G$ is max-stable, then there exist functions $a(s) > 0, b(s) \in \mathbb{R}, s \in \mathbb{R}^+$, such that

\[
G'(a(s)x + b(s)) = G(x).
\]

**Proof.** This follows essentially by interpolation. We have that

\[
G^{[ns]}(a_{[ns]}x + b_{[ns]}) = G(x).
\]

But

\[
G^n(a_{[ns]}x + b_{[ns]}) = G^{[ns]/s}(a_{[ns]}x + b_{[ns]})G^{n-[ns]/s}(a_{[ns]}x + b_{[ns]})
\]

\[
= G^{1/s}(x)G^{n-[ns]/s}(a_{[ns]}x + b_{[ns]}).
\]

As $n \uparrow \infty$, the last factor tends to one (as the exponent remains bounded), and so

\[
G^n(a_{[ns]}x + b_{[ns]}) \xrightarrow{w} G^{1/s}(x),
\]

and

\[
G^n(a_nx + b_n) \xrightarrow{w} G(x).
\]

Thus by Khintchine’s theorem,

\[
a_{[ns]}/a_n \to a(s), \quad (b_n - b_{[ns]})/a_n \to b(s),
\]

and

\[
G^{1/s}(x) = G(a(s)x + b(s)).
\]

\qed
1.2.4 The extremal types theorem.

Definition 1.13. Two distribution functions, \( G, H \), are called “of the same type”, if and only if there exists \( a > 0, b \in \mathbb{R} \) such that
\[
G(x) = H(ax + b). \tag{1.2.56}
\]

We have seen that the only distributions that can occur as extremal distributions are max-stable distributions. We will now classify these distributions.

Theorem 1.14. Any max-stable distribution is of the same type as one of the following three distributions:

(I) The Gumbel distribution,
\[
G(x) = e^{-e^{-x}}. \tag{1.2.57}
\]

(II) The Fréchet distribution with parameter \( \alpha > 0 \),
\[
G(x) = \begin{cases} 
0, & \text{if } x \leq 0 \\
 e^{-\frac{x}{\alpha}}, & \text{if } x > 0.
\end{cases} \tag{1.2.58}
\]

(III) The Weibull distribution with parameter \( \alpha > 0 \),
\[
G(x) = \begin{cases} 
 e^{(-x)^{\alpha}}, & \text{if } x < 0 \\
1, & \text{if } x \geq 0.
\end{cases} \tag{1.2.59}
\]

We do not give the rather lengthy proof of this theorem. It can be found in [57].

Let us state as an immediate corollary the so-called extremal types theorem.

Theorem 1.15. Let \( X_i, i \in \mathbb{N} \) be a sequence of i.i.d. random variables. If there exist sequences \( a_n > 0, b_n \in \mathbb{R} \), and a non-degenerate probability distribution function, \( G \), such that
\[
\mathbb{P}(a_n(M_n - b_n) \leq x) \underset{n \to \infty}{\to} G(x), \tag{1.2.60}
\]
then \( G(x) \) is of the same type as one of the three extremal-type distributions.

Note that it is not true, of course, that for arbitrary distributions of the variables \( X_i \) it is possible to obtain a nontrivial limit as in (1.2.60).

The following theorem gives necessary and sufficient conditions. We set \( x_F \equiv \sup \{x : F(x) < 1\} \).

Theorem 1.16. The following conditions are necessary and sufficient for a distribution function, \( F \), to belong to the domain of attraction of the three extremal types:

Fréchet: \( x_F = +\infty \),
\[
\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}, \quad \forall x > 0, \alpha > 0 \tag{1.2.61}
\]
1.3 Level-crossings and the distribution of the \( k \)-th maxima.

Weibull: \( x_F = < +\infty \),

\[
\lim_{h \to 0} \frac{1 - F(x_F - xh)}{1 - F(x_F - h)} = x^\alpha, \quad \forall x > 0, \quad \alpha > 0 \tag{1.2.62}
\]

Gumbel: \( \exists g(t) > 0 \),

\[
\lim_{t \to x} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-t}, \quad \forall t. \tag{1.2.63}
\]

This theorem is not of too much interest for us later and we refer to [65] for the proof.

1.3 Level-crossings and the distribution of the \( k \)-th maxima.

In the previous section we have answered the question of the distribution of the maximum of \( n \) iid random variables. It is natural to ask for more, i.e. for the joint distribution of the maximum, the second largest, third largest, etc.

From what we have seen, the levels \( u_n \) for which \( P(X_n > u_n) \sim \tau/n \) will play a crucial rôle. A natural variable to study is \( M_n^k \), the value of the \( k \)-th largest of the first \( n \) variables \( X_i \).

It will be useful to introduce here the notion of order statistics.

Definition 1.17. Let \( X_1, \ldots, X_n \) be real numbers. Then we denote by \( M_n^1, \ldots, M_n^n \) its order statistics, i.e. for some permutation, \( \pi \), of \( n \) numbers, \( M_n^k = X_{\pi(k)} \), and

\[
M_n^n \leq M_n^{n-1} \leq \cdots \leq M_n^2 \leq M_n^1 = M_n. \tag{1.3.1}
\]

We will also introduce the notation

\[
S_n(u) \equiv \# \{ i \leq n : X_i > u \} \tag{1.3.2}
\]

for the number of exceedences of the level \( u \). Obviously we have the relation

\[
P(M_n^k \leq u) = P(S_n(u) < k). \tag{1.3.3}
\]

The following result states that the number of exceedances of an extremal level \( u_n \) is Poisson distributed.

Theorem 1.18. Let \( X_i \) be iid random variables with common distribution \( F \). If \( u_n \) is such that

\[
n(1 - F(u_n)) \to \tau, \quad 0 < \tau < \infty, \tag{1.3.4}
\]

then

\[
P(M_n^k \leq u_n) = P(S_n(u_n) < k) \to e^{-\tau} \sum_{s=0}^{k-1} \frac{\tau^s}{s!} \tag{1.3.5}
\]
Proof. The proof of this lemma is quite simple. We just need to consider all possible ways to realise the event \( \{ S_n(u_n) = s \} \). Namely
\[
\Pr \{ S_n(u_n) = s \} = \sum \prod_{i=1}^{\ell} \Pr \{ X_i > u_n \} \prod_{j \notin \{1, \ldots, \ell\}} \Pr \{ X_j \leq u_n \}
= \binom{n}{s} (1 - F(u_n))^s F(u_n)^{n-s}
= \frac{1}{s! n^s (n-s)!} \left[ n(1 - F(u_n)) \right]^s \left[ F^n(u_n) \right]^{1-s/n}.
\]

But, for any \( s \) fixed, \( n(1 - F(u_n)) \to \tau, F^n(u_n) \to e^{-\tau}, s/n \to 0, \) and \( \frac{n^s}{n^s (n-s)!} \to 1 \). Thus
\[
\Pr \{ S_n(u_n) = s \} \to \frac{\tau^s}{s!} e^{-\tau}.
\]

Summing over all \( s < k \) gives the assertion of the theorem. □

Using very much the same sort of reasoning, one can generalise the question answered above to that of the numbers of exceedances of several extremal levels.

**Theorem 1.19.** Let \( u_1 > u_2 > \cdots > u_r \) such that
\[
n(1 - F(u_\ell)) \to \tau_\ell,
\]
with
\[
0 < \tau_1 < \tau_2 < \cdots < \tau_r < \infty.
\]

Then, under the assumptions of the preceding theorem, with \( S_n = S_n(u_\ell) \),
\[
\Pr \{ S_n^1 = k_1, S_n^2 = k_2, \ldots, S_n^r = k_r \} \to \frac{k_1!}{k_1!} \frac{(\tau_2 - \tau_1)^{k_2}}{k_2!} \cdots \frac{(\tau_r - \tau_{r-1})^{k_r}}{k_r!} e^{-\tau_r}.
\]

Proof. Again, we just have to count the number of arrangements that will place the desired number of variables in the respective intervals. Then
\[
\Pr \{ S_n^1 = k_1, S_n^2 = k_2, \ldots, S_n^r = k_r \} = \binom{n}{k_1, \ldots, k_r} \prod \Pr \{ X_{k_1+1} \geq X_{k_1+2}, \ldots, X_{k_2+k_1} > u_2, \ldots, u_1, \ldots u_{k_1-1} \geq X_{k_1+\cdots+k_2}, \ldots, u_1, \ldots, u_{k_r-1} \geq X_{k_1+\cdots+k_r}, \ldots, u_1 \}
= \binom{n}{k_1, \ldots, k_r} (1 - F(u_1)^{k_1}) F(u_1)^{k_1} \cdots (1 - F(u_n)^{k_r}) F(u_n)^{k_r}
\times F^{n-k_1-\cdots-k_r}(u_r).
\]
1.3 Level-crossings and the distribution of the $k$-th maxima.

\[ \frac{1}{n} \left[ n(1 - F(u_{n}^{'l-1})) - n(1 - F(u_{n}^{l-1})) \right] \]

and use that \( n(1 - F(u_{n}^{l-1})) \to \tau_{l} - \tau_{l-1} \). Proceeding otherwise as in the proof of Theorem 1.18, we arrive at (1.3.10) \( \square \)
Chapter 2
Extremal processes

In this chapter we develop and complete the description of the collection of “extremal values” of a stochastic process. Here we will develop this theory in the language of point processes. We begin with some background on this subject and the particular class of processes that will turn out to be fundamental, the Poisson point processes. For more details, see [65, 50, 30, 52].

2.1 Point processes

Point processes are designed to describe the probabilistic structure of point sets in some metric space, for our purposes $\mathbb{R}^d$. For reasons that may not be obvious immediately, a convenient way to represent a collection of points $x_i$ in $\mathbb{R}^d$ is by associating to them a point measure.

Let us first consider a single point $x$. We consider the usual Borel-sigma algebra, $\mathcal{B} = \mathcal{B}(\mathbb{R}^d)$, of $\mathbb{R}^d$, that is generated from the open sets in the open sets in the Euclidean topology of $\mathbb{R}^d$. Given $x \in \mathbb{R}^d$, we define the Dirac measure, $\delta_x$, such that, for any Borel set $A \in \mathcal{B}$,

$$\delta_x(A) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A. \end{cases} \tag{2.1.1}$$

**Definition 2.1.** A point measure on $\mathbb{R}^d$ is a measure, $\mu$, on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, such that there exists a countable collection of points, $\{x_i \in \mathbb{R}^d, i \in \mathbb{N}\}$, such that

$$\mu = \sum_{i=1}^{\infty} \delta_{x_i}, \tag{2.1.2}$$

and if $K$ is compact, then $\mu(K) < \infty$.  

15
Note that the points $x_i$ need not be all distinct. The set $S_\mu = \{ x \in \mathbb{R}^d : \mu(x) \neq 0 \}$ is called the support of $\mu$. A point measure such that for all $x \in \mathbb{R}^d$, $\mu(x) \leq 1$ is called simple.

We denote by $M_p(\mathbb{R}^d)$ the set of all point measures on $\mathbb{R}^d$. We equip this set with the sigma-algebra $\mathcal{M}_p(\mathbb{R}^d)$, the smallest sigma algebra that contains all subsets of $M_p(\mathbb{R}^d)$ of the form $\{ \mu \in M_m(\mathbb{R}^d) : \mu(F) \in B \}$, where $F \in \mathcal{B}(\mathbb{R}^d)$ and $B \in \mathcal{B}([0, \infty))$. $\mathcal{M}_p(\mathbb{R}^d)$ is also characterised by saying that it is the smallest sigma-algebra that makes the evaluation maps, $\mu \mapsto \mu(F)$, measurable for all Borel sets $F \in \mathcal{B}(\mathbb{R}^d)$.

**Definition 2.2.** A point process, $N$, is a random variable taking values in $M_p(\mathbb{R}^d)$, i.e. a measurable map, $N : (\Omega, \mathcal{F}, \mathbb{P}) \to M_p(\mathbb{R}^d)$, from a probability space to the space of point measures.

This looks very fancy, but in reality things are quite down-to-earth:

**Proposition 2.3.** A map $N : \Omega \to M_p(\mathbb{R}^d)$ is a point process, if and only if the map $N(\cdot, F) : \omega \to N(\omega, F)$, is measurable from $(\Omega, \mathcal{F}) \to ([0, \infty], \mathcal{B}([0, \infty)))$, for any Borel set $F$, i.e. if $N(F)$ is a real random variable.

**Proof.** Let us first prove necessity, which should be obvious. In fact, since $\omega \to N(\omega, \cdot)$ is measurable into $(M_p(\mathbb{R}^d), \mathcal{M}_p(\mathbb{R}^d))$, and $\mu \mapsto \mu(F)$ is measurable from this space into $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, the composition of these maps is also measurable.

Next we prove sufficiency. Define the set

$$\mathcal{G} \equiv \{ A \in \mathcal{M}_p(\mathbb{R}^d) : N^{-1}A \in \mathcal{F} \}. \quad (2.1.3)$$

This set is a sigma-algebra and $N$ is measurable from $(\Omega, \mathcal{F}) \to (M_p(\mathbb{R}^d), \mathcal{G})$ by definition. But $\mathcal{G}$ contains all sets of the form $\{ \mu \in M_p(\mathbb{R}^d) : \mu(F) \in B \}$, since

$$N^{-1}\{ \mu \in M_p(\mathbb{R}^d) : \mu(F) \in B \} = \{ \omega \in \Omega : N(\omega, F) \in B \} \in \mathcal{F}, \quad (2.1.4)$$

since $N(\cdot, F)$ is measurable. Thus $\mathcal{G} \supset \mathcal{M}_p(\mathbb{R}^d)$, and $N$ is measurable a fortiori as a map from the smaller sigma-algebra. $\square$

We will have need to find criteria for convergence of point processes. For this we recall some basic notions from standard measure theory. If $\mathcal{B}$ is a Borel-sigma algebra, of a metric space $E$, then $\mathcal{T} \subset \mathcal{B}$ is called a $\Pi$-system, if $\mathcal{T}$ is closed under finite intersections; $\mathcal{G} \subset \mathcal{B}$ is called a $\lambda$-system, or a sigma-additive class, if

(i) $E \in \mathcal{G}$,

(ii) If $A, B \in \mathcal{G}$, and $A \supset B$, then $A \setminus B \in \mathcal{G}$,

(iii) If $A_n \in \mathcal{G}$ and $A_n \subset A_{n+1}$, then $\lim_{n \to \infty} A_n \in \mathcal{G}$.

The following useful observation is called Dynkin’s theorem.

**Theorem 2.4.** If $\mathcal{T}$ is a $\Pi$-system and $\mathcal{G}$ is a $\lambda$-system, then $\mathcal{G} \supset \mathcal{T}$ implies that $\mathcal{G}$ contains the smallest sigma-algebra containing $\mathcal{T}$. 

2.1 Point processes

The most useful application of Dynkin’s theorem is the observation that, if two probability measures are equal on a \( \Pi \)-system that generates the sigma-algebra, then they are equal on the sigma-algebra. (since the set on which the two measures coincide forms a \( \lambda \)-system containing \( \mathcal{T} \)).

As a consequence we can further restrict the criteria to be verified for \( N \) to be a point process. In particular, we can restrict the class of \( F \)’s for which \( N(\cdot,F) \) need to be measurable to bounded rectangles.

**Proposition 2.5.** Suppose that \( \mathcal{T} \) are relatively compact sets in \( \mathcal{B} \) satisfying

(i) \( \mathcal{T} \) is a \( \Pi \)-system,
(ii) The smallest sigma-algebra containing \( \mathcal{T} \) is \( \mathcal{B} \),
(iii) Either, there exists \( E_n \in \mathcal{T} \), such that \( E_n \uparrow E \), or there exists a partition, \( \{E_n\} \), of \( E \) with \( \cup_n E_n = E \), with \( E_n \subset \mathcal{T} \).

Then \( N \) is a point process on \( (\Omega, \mathcal{F}) \) in \( (E, \mathcal{B}) \), if and only if the map \( N(\cdot,I) : \omega \to N(\omega,I) \) is measurable for any \( I \in \mathcal{T} \).

**Exercise.** Check that the set of all finite collections bounded (semi-open) rectangles forms indeed a \( \Pi \)-system for \( E = \mathbb{R}^d \) that satisfies the hypotheses of the proposition.

**Corollary 2.6.** Let \( \mathcal{T} \) satisfy the hypothesis of Proposition 2.5 and set

\[
\mathcal{G} \equiv \{ \mu : \mu(I_j) = n_j, 1 \leq j \leq k \}, \quad k \in \mathbb{N}, I_j \in \mathcal{T}, n_j \geq 0 \}. \tag{2.1.5}
\]

Then the smallest sigma-algebra containing \( \mathcal{G} \) is \( \mathcal{M}_p(\mathbb{R}^d) \) and \( \mathcal{G} \) is a \( \Pi \)-system.

Next we show that the law, \( P_N \), of a point process is determined by the law of the collections of random variables \( N(F_n), F_n \in \mathcal{B}(\mathbb{R}^d) \).

**Proposition 2.7.** Let \( N \) be a point process in \( (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \) and suppose that \( \mathcal{T} \) is as in Proposition 2.5. Define the mass functions

\[
P_{I_1,...,I_k}(n_1,\ldots,n_k) \equiv \mathbb{P}(N(I_j) = n_j, \forall 1 \leq j \leq k), \tag{2.1.6}
\]

for \( I_j \in \mathcal{T}, n_j \geq 0 \). Then \( P_N \) is uniquely determined by the collection

\[
\{P_{I_1,...,I_k}, k \in \mathbb{N}, I_j \in \mathcal{T}\}. \tag{2.1.7}
\]

We need some further notions.

**Definition 2.8.** Two point processes, \( N_1, N_2 \) are are independent, if and only if, for any collection \( F_j \in \mathcal{B}, G_j \in \mathcal{B} \), the vectors

\[
(N_1(F_j), 1 \leq j \leq k) \quad \text{and} \quad (N_2(G_j), 1 \leq j \leq \ell) \tag{2.1.8}
\]

are independent random vectors.
Definition 2.9. The set function \( \lambda \) defined on \( (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \) defined by
\[
\lambda(F) \equiv \mathbb{E}N(F) = \int_{M_p(\mathbb{R}^d)} \mu(F)P_N(d\mu),
\]
for \( F \in \mathcal{B} \), is a measure, called the intensity measure of the point process \( N \).

For measurable functions \( f : \mathbb{R}^d \to \mathbb{R}_+ \), we define
\[
N(\omega, f) \equiv \int_{\mathbb{R}^d} f(x)N(\omega, dx).
\]
Then \( N(\cdot, f) \) is a random variable. We have that
\[
\mathbb{E}N(f) = \lambda(f) = \int_{\mathbb{R}^d} f(x)\lambda(dx).
\]

2.2 Laplace functionals

If \( Q \) is a probability measure on \( (M_p, M_p) \), the Laplace transform of \( Q \) is a map, \( \psi \) from non-negative Borel functions on \( \mathbb{R}^d \) to \( \mathbb{R}_+ \), defined as
\[
\psi(f) \equiv \int_{M_p} \exp \left( -\int_{\mathbb{R}^d} f(x)\mu(dx) \right) Q(d\mu).
\]

If \( N \) is a point process, the Laplace functional of \( N \) is
\[
\psi_N(f) \equiv \mathbb{E}\exp\left( -N(\omega, f) \right) = \int_{M_p} \exp\left( -\int_{\mathbb{R}^d} f(x)\mu(dx) \right) P_N(d\mu).
\]

Proposition 2.10. The Laplace functional, \( \psi_N \), of a point process, \( N \), determines \( N \) uniquely.

Proof. For \( k \geq 1 \), and \( F_1, \ldots, F_k \in \mathcal{B}, c_1, \ldots, c_k \geq 0 \), let \( f = \sum_{i=1}^k c_i 1_{F_i}(x) \). Then
\[
N(\omega, f) = \sum_{i=1}^k c_i N(\omega, F_i),
\]
and
\[
\psi_N(f) = \mathbb{E}\exp\left( -\sum_{i=1}^k c_i N(F_i) \right).
\]
This is the Laplace transform of the vector \( (N(F_i), 1 \leq i \leq k) \), that determines uniquely its law. Hence the proposition follows from Proposition 2.7 □
2.3 Poisson point processes.

One can restrict the set of functions that is required for the Laplace functionals to determine the measure considerably, as one can see from the proof above.

2.3 Poisson point processes.

The most important class of point processes for our purposes will be Poisson point processes.

**Definition 2.11.** Let $\lambda$ be a $\sigma$-finite, positive measure on $\mathbb{R}^d$. Then a point process, $N$, is called a Poisson point process with intensity measure $\lambda$ (PPP($\lambda$)), if

(i) For any $F \in \mathcal{B}(\mathbb{R}^d)$, and $k \in \mathbb{N}$,
$$
\mathbb{P}(N(F) = k) = \begin{cases} 
  e^{-\lambda(F)} \frac{(\lambda(F))^k}{k!}, & \text{if } \lambda(F) < \infty, \\
  0, & \text{if } \lambda(F) = \infty.
\end{cases} 
$$

(ii) If $F, G \in \mathcal{B}$ are disjoint sets, then $N(F)$ and $N(G)$ are independent random variables.

In the next theorem we will assert the existence of a Poisson point process with any desired intensity measure. In the proof we will give an explicit construction of such a process.

**Proposition 2.12.** (i) PPP($\lambda$) exists, and its law is uniquely determined by the requirements of the definition.

(ii) The Laplace functional of PPP($\lambda$) is given, for $f \geq 0$, by
$$
\Psi_N(f) = \exp \left(-\int_{\mathbb{R}^d} (1 - e^{-f(x)}) \lambda(dx) \right). 
$$

**Proof.** Since we know that the Laplace functional determines a point process, in order to prove that the conditions of the definition uniquely determine the PPP($\lambda$), we show that they determine the form (2.3.2) of the Laplace functional. Thus suppose that $N$ is a PPP($\lambda$). Let $f = c I_F$. Then
$$
\Psi_N(f) = \mathbb{E}[\exp\{-N(f)\}] = \mathbb{E}[\exp\{-cN(F)\}] 
= \sum_{k=0}^{\infty} e^{-ck} e^{-\lambda(F)} \frac{(\lambda(F))^k}{k!} = e^{(e^c-1)\lambda(F)} 
= \exp\left(-\int_{\mathbb{R}^d} (1 - e^{-f(x)}) \lambda(dx) \right),
$$
which is the desired form. Next, if $F_i$ are disjoint, and $f = \sum_{i=1}^{k} c_i I_{F_i}$, it is straightforward to see that
\[ \Psi_N(f) = \mathbb{E}\left[ \exp\left( -\sum_{i=1}^{k} c_i N(F_i) \right) \right] = \prod_{i=1}^{k} \mathbb{E}[\exp(-c_i N(F_i))], \quad (2.3.4) \]
due to the independence assumption (ii); a simple calculation shows that this yields again the desired form. Finally, for general \( f \), we can choose a sequence, \( f_n \), of the form considered, such that \( f_n \uparrow f \). By monotone convergence then \( N(f_n) \uparrow N(f) \). On the other hand, since \( e^{-N(g)} \leq 1 \), we get from dominated convergence that \( \Psi_N(f_n) = \mathbb{E}e^{-N(f_n)} \uparrow \mathbb{E}e^{-N(f)} = \Psi_N(f) \). \quad (2.3.5)

But, since \( 1 - e^{-f_n(x)} \uparrow 1 - e^{-f(x)} \), and monotone convergence gives once more \( \Psi_N(f_n) = \exp\left( \int (1 - e^{-f_n(x)}) \lambda(dx) \right) \uparrow \exp\left( \int (1 - e^{-f(x)}) \lambda(dx) \right). \) \quad (2.3.6)

On the other hand, given the form of the Laplace functional, it is trivial to verify that the conditions of the definition hold, by choosing suitable functions \( f \).

Finally we turn to the construction of \( \text{PPP}(\lambda) \). Let us first consider the case \( \lambda(\mathbb{R}^d) < \infty \). Then construct

(i) A Poisson random variable, \( \tau \), of parameter \( \lambda(\mathbb{R}^d) \).

(ii) A family, \( X_i, i \in \mathbb{N} \), of independent, \( \mathbb{R}^d \)-valued random variables with common distribution \( \lambda \). This family is independent of \( \tau \).

Then set \( N^* \equiv \sum_{i=1}^{\tau} \delta_{X_i} \). \quad (2.3.7)

The easiest way to verify that \( N^* \) is a \( \text{PPP}(\lambda) \) is to compute its Laplace functional. This is left as an easy exercise.

To deal with the case when \( \lambda(\mathbb{R}^d) \) is infinite, decompose \( \lambda \) into a countable sum of finite measures, \( \lambda_k \), that are just the restriction of \( \lambda \) to a finite set \( F_k \), where the \( F_k \) form a partition of \( \mathbb{R}^d \). Then \( N^* \) is just the sum of independent \( \text{PPP}(\lambda_k) \) \( N_k^* \).

### 2.4 Convergence of point processes

Before we turn to applications to extremal processes, we still have to discuss the notion of convergence of point processes. As point processes are probability distributions on the space of point measures, we will naturally think about weak convergence. This means that we will say that a sequence of point processes, \( N_n \), converges weakly to a point process, \( N \), if for all continuous functions, \( f \), on the space of point measures,

\[ \mathbb{E}f(N_n) \rightarrow \mathbb{E}f(N). \] \quad (2.4.1)

However, to understand what this means, we must discuss what continuous functions on the space of point measures are, i.e. we must introduce a topology on the...
set of point measures. The appropriate topology for our purposes will be that of vague convergence.

2.4 Vague convergence.

We consider the space $\mathbb{R}^d$ equipped with its natural Euclidean metric. Clearly $\mathbb{R}^d$ is a complete, separable metric space. We will denote by $C_0(\mathbb{R}^d)$ the set of continuous real-valued functions on $\mathbb{R}^d$ that have compact support; $C_0^+(\mathbb{R}^n)$ denotes the subset of non-negative functions. We consider $\mathcal{M}_+(\mathbb{R}^d)$ the set of all $\sigma$-finite, positive measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. We denote by $\mathcal{M}_0^+(\mathbb{R}^d)$ the smallest sigma-algebra of subsets of $\mathcal{M}_+(\mathbb{R}^d)$ that makes the maps $m \rightarrow m(f)$ measurable for all $f \in C_0^+(\mathbb{R}^d)$.

Definition 2.13. A sequence of measures, $\mu_n \in \mathcal{M}_+(\mathbb{R}^d)$ converges vaguely to a measure $\mu \in \mathcal{M}_+(\mathbb{R}^d)$, if, for all $f \in C_0^+(\mathbb{R}^d)$,

$$
\mu_n(f) \rightarrow \mu(f).
$$

Vague convergence defines a topology on the space of measures. Typical open neighbourhoods are of the form

$$
B_{\epsilon_1, \ldots, \epsilon_k}(\mu) \equiv \{ \nu \in \mathcal{M}_+(\mathbb{R}^d) : \forall i=1 \ldots k |\nu(f_i) - \mu(f_i)| < \epsilon_i \},
$$

i.e. to test the closeness of two measures, we test it on their expectations on finite collections of continuous, positive functions with compact support. Given this topology, on can of course define the corresponding Borel sigma algebra, $\mathcal{B}(\mathcal{M}_+(\mathbb{R}^d))$, which (fortunately) turns out to coincide with the sigma algebra $\mathcal{M}_+(\mathbb{R}^d)$ introduced before.

The following properties of vague convergence are useful.

Proposition 2.14. Let $\mu_n, n \in \mathbb{N}$ be in $\mathcal{M}_+(\mathbb{R}^d)$. Then the following statements are equivalent:

(i) $\mu_n$ converges vaguely to $\mu$, $\mu_n \xrightarrow{v} \mu$.
(ii) $\mu_n(B) \rightarrow \mu(B)$ for all relatively compact sets, $B$, such that $\mu(\partial B) = 0$.
(iii) $\limsup_{n \rightarrow \infty} \mu_n(K) \leq \mu(K)$ and $\liminf_{n \rightarrow \infty} \mu_n(G) \geq \mu(G)$, for all compact $K$, and all open, relatively compact $G$.

In the case of point measures, we would of course like to see that the point where the sequence of vaguely convergent measures are located converge. The following proposition tells us that this is true.

Proposition 2.15. Let $\mu_n, n \in \mathbb{N}$, and $\mu$ be in $\mathcal{M}_p(\mathbb{R}^d)$, and $\mu_n \xrightarrow{v} \mu$. Let $K$ be a compact set with $\mu(\partial K) = 0$. Then we have a labelling of the points of $\mu_n$, for $n \geq n(K)$ large enough, such that
such that $(x^n_1, \ldots, x^n_p) \to (x_1, \ldots, x_p)$.

Another useful and unsurprising fact is that

**Proposition 2.16.** The set $M_p(\mathbb{R}^d)$ is vaguely closed in $M_+(\mathbb{R}^d)$.

Thus, in particular, the limit of a sequence of point measures, will, if it exists as a $\sigma$-finite measure, be again a point measure.

Finally, we need a criterion that describes relatively compact sets.

**Proposition 2.17 (Proposition 3.16 in [65]).** A subset $M$ of $M_+(E)$ or $M_p(E)$ is relatively compact, if and only if one of the following holds:

- For all $f \in C_0^+(E)$, 
  \[ \sup_{\mu \in M} \mu(f) < \infty, \]  
  (2.4.5)

- For all relatively compact $B \in \mathcal{B}(E)$, 
  \[ \sup_{\mu \in M} \mu(B) < \infty, \]  
  (2.4.6)

For the proof, see [65].

**Proposition 2.18.** The topology of vague convergence can be metrised and turns $M_+$ into a complete, separable metric space.

Although we will not use the corresponding metric directly, it may be nice to see how this can be constructed. We therefore give a proof of the proposition that constructs such a metric.

**Proof.** The idea is to first find a countable collection of functions, $h_i \in C_0^+(\mathbb{R}^d)$, such that $\mu_n \rightharpoonup \mu$ if and only if, for all $i \in \mathbb{N}$, $\mu_n(h_i) \to \mu(h_i)$. The construction below is from [50]. Take a family $G_i, i \in \mathbb{N}$, that form a base of relatively compact sets, and assume it to be closed under finite unions and finite intersections. One can find (by Uryson’s theorem), families of functions $f_{i,n}, g_{i,n} \in C_0^+$, such that

\[ f_{i,n} \uparrow 1_{G_i}, \quad g_{i,n} \downarrow 1_{G_i}, \]  
(2.4.7)

Take the countable set of functions $g_{i,n}, f_{i,n}$ as the collection $h_i$. Now $\mu \in M_+$ is determined by its values on the $h_i$. For, first of all, $\mu(G_i)$ is determined by these values, since

\[ \mu(f_{i,n}) \uparrow \mu(G_i) \quad \text{and} \quad \mu(g_{i,n}) \downarrow \mu(G_i). \]  
(2.4.8)

But the family $G_i$ is a $\Pi$-system that generates the sigma-algebra $\mathcal{B}(\mathbb{R}^d)$, and so the values $\mu(G_i)$ determine $\mu$.

Now, $\mu_n \rightharpoonup \mu$ if and only if, for all $h_i$, $\mu_n(h_i) \to c_i = \mu(h_i)$. 

\[ \mu_n \rightharpoonup \mu \quad \text{iff} \quad \mu_n \rightharpoonup \mu. \]
2.4 Convergence of point processes

From here the idea is simple: Define
\[ d(\mu, \nu) \equiv \sum_{i=1}^{\infty} 2^{-i} \left( 1 - e^{-|\mu(h_i) - \nu(h_i)|} \right) \]  \hspace{1cm} (2.4.9)

Indeed, if \( D(\mu_n, \mu) \downarrow 0 \), then for each \( \ell \), \( |\mu_n(h) - \mu(h)| \downarrow 0 \), and conversely. \( \square \)

It is not very difficult to verify that this metric is complete and separable.

### 2.4.2 Weak convergence.

Having established the space of \( \sigma \)-finite measures as a complete, separable metric space, we can think of weak convergence of probability measures on this space just as if we were working on an Euclidean space.

One very useful fact about weak convergence is Skorokhod’s theorem, that relates weak convergence to \textit{almost sure convergence}.

**Theorem 2.19.** Let \( X_n, n = 0, 1, \ldots \) be a sequence of random variables on a complete separable metric space. Then \( X_n \) converges weakly to a random variable \( X_0 \), iff and only if there exists a family of random variables \( X_n^* \), defined on the probability space \(([0, 1], \mathcal{B}([0, 1]), m)\), where \( m \) is the Lebesgue measure, such that

(i) For each \( n \), \( X_n \overset{D}{=} X_n^* \), and
(ii) \( X_n^* \to X_0^* \), almost surely.

(for a proof, see [12]). While weak convergence usually means that the actual realisation of the sequence of random variables do not converge at all and oscillate widely, Skorokhod’s theorem says that it is possible to find an “equally likely” sequence of random variables, \( X_n^* \), that do themselves converge, with probability one.

Such a construction is easy in the case when the random variables take values in \( \mathbb{R} \).

In that case, we associate with the random variable \( X_n \) (whose distribution function is \( F_n \)), that for simplicity we may assume strictly increasing), the random variable \( X_n^*(t) \equiv F_n^{-1}(t) \). It is easy to see that

\[ m(X_n^* \leq x) = \int_0^1 1_{F_n^{-1}(t) \leq x} \, dt = F_n(x) = P(X_n \leq x). \]  \hspace{1cm} (2.4.10)

On the other hand, if \( P(X_n \leq x) \to P(X_0 \leq x) \), for all points of continuity of \( F_0 \), that means that for Lebesgue almost all \( t \), \( F_n^{-1}(t) \to F_0^{-1}(t) \), i.e. \( X_n^* \to X_0^* \), \( m \)-almost surely.

Skorokhod’s theorem is very useful to extract important consequences from weak convergence. In particular, it allows to prove the convergence of certain functionals of sequences of weakly convergent random variables, which otherwise would not be obvious.

When proving weak convergence, one usually does this in two steps.
• Prove tightness of the sequence of laws.
• Prove convergence of finite dimensional distributions.

A useful tightness criterion is the following (see [65], Lemma 3.20).

**Lemma 2.20.** A sequence of point processes \( \xi_n \) is tight, iff for any \( f \in C_0^+ (E) \), the sequence \( \xi_n f \) is tight as a sequence of real random variables.

**Proof.** Assume that all \( f \in C_0^+ (E) \), the sequence \( \xi_n f \) is tight. Let \( g_i \in C_0^+ (E) \) be a sequence of functions such that \( g_i \uparrow 1 \), as \( i \uparrow 1 \). Then for any \( \epsilon > 0 \), there exists \( c_i < \infty \), such that

\[
P(\xi_n (g_i) > c_i) \leq \epsilon 2^{-i}.
\]

(2.4.11)

Let \( M \equiv \cap_{i \geq 1} \{ \mu \in M_+(E) : \mu (g_i) \leq c_i \} \). We claim that this implies that for any \( f \in C_0^+ (E) \), we have that \( \sup_{\mu \in M} \mu (f) < \infty \). To see this, since \( g_i \uparrow 1 \), for any \( f \) there will be \( K_0 < \infty \) and \( i_0 < \infty \), such that \( f \leq K g_{i_0} \), and hence

\[
\sup_{\mu \in M} \mu (f) \leq \sup_{\mu \in M} K_0 \mu (g_{i_0}) \leq K_0 c_{i_0}.
\]

(2.4.12)

This implies that \( M \) is relatively compact in \( M_+(E) \). Finally,

\[
P (\xi_n \not\in M) \leq P (\xi \not\in M) \leq \sum_{i \geq 1} P (\xi_n (g_i) > c_i) \leq 2\epsilon.
\]

(2.4.13)

Hence, for any \( \epsilon > 0 \), there is a compact set such that for all \( n \), \( \xi_n \) is in this set with probability at least \( 1 - \epsilon \). This implies that \( \xi_n \) is tight. The converse statement is trivial. \( \Box \)

**Corollary 2.21.** Let \( E = \mathbb{R} \). Assume that for all \( a \) the sequence \( \xi_n (1_{x > a}) \) is tight. Then \( \xi_n \) is tight.

**Proof.** Obviously, we can chose the sequence \( g_i \) in the proof above as \( g_i (x) = 1_{x > a_i} \) for \( a_i \uparrow \infty \). \( \Box \)

Clearly, Laplace functionals are a good tool to verify weak convergence.

**Proposition 2.22.** Let \( N_n \) be a sequence of point processes. Then \( N_n \) converges weakly to \( N \), if and only if all Laplace functionals converge, more precisely, if for all \( f \in C_0^+ (E) \),

\[
\Psi_{N_n} (f) \to \Psi_N (f).
\]

(2.4.14)

**Proof.** We just sketch this. (2.4.14) asserts that the Laplace transforms of the positive random variables \( N_n (f) \) converge to a random variables, which implies convergence in distribution and hence tightness. On the other hand, \( \Psi_N (f) \) determines the law of \( N \), and hence there can be only one limit point of the sequences \( N_n \), which implies weak convergence. The converse assertion is trivial. \( \Box \)
2.4 Convergence of point processes

Of course we do not really need to check convergence for all \( f \in C_0^+ \). For instance, in the case \( E = \mathbb{R} \), we may choose the class of functions of the form \( f(x) = \sum_{i=1}^{k} c_i \mathbb{1}_{x > u_i}, \) \( k \in \mathbb{N}, c_i > 0, u_i \in \mathbb{R} \). Clearly, the Laplace functionals evaluated on these functions determine the Laplace transforms of the vectors \( \{N((u_1, \infty)), \ldots, N_k((u_k, \infty))\} \), and hence the probabilities (assuming the \( u_i \) are an increasing sequence)

\[
P(N((u_1, \infty))) = m_1, \ldots, N_k((u_k, \infty)) = m_k
\]

(2.4.15)

and hence the mass functions, that we know to determine the law of \( N \) by Proposition 2.7.

Another useful criterion for weak convergence of point processes is provided by Kallenberg’s theorem [30].

Theorem 2.23. Assume that \( \xi \) is a simple point process on a metric space \( E \), and \( \mathcal{T} \) is a \( \Pi \)-system of relatively compact open sets, and that for \( I \in \mathcal{T} \),

\[
P(\xi(\partial I) = 0) = 1.
\]

(2.4.16)

If \( \xi_n, n \in \mathbb{N} \) are point processes on \( E \), and for all \( I \in \mathcal{T} \),

\[
\lim_{n \to \infty} P(\xi_n(I) = 0) = P(\xi(I) = 0),
\]

(2.4.17)

and

\[
\lim_{n \to \infty} E\xi_n(I) = E\xi(I) < \infty,
\]

(2.4.18)

then

\[
\xi_n \xrightarrow{w} \xi.
\]

(2.4.19)

Remark 2.24. The \( \Pi \)-system, \( \mathcal{T} \), can be chosen, in the case \( E = \mathbb{R}^d \), as finite unions of bounded rectangles.

Proof. The key observation needed to prove the theorem is that simple point processes are uniquely determined by their avoidance function. This seems rather intuitive, in particular in the case \( E = \mathbb{R} \): if we know the probability that in an interval there is no point, we know the distribution of the gape between points, and thus the distribution of the points.

Let us note that we can write a point measure, \( \mu \), as

\[
\mu = \sum_{y \in S} c_y \delta_y,
\]

(2.4.20)

where \( S \) is the support of the point measure and \( c_y \) are integers. We can associate to \( \mu \) the simple point measure

\[
T^* \mu = \mu^* = \sum_{y \in S} \delta_y,
\]

(2.4.21)
Then it is true that the map $T^*$ is measurable, and that, if $\xi_1$ and $\xi_2$ are point measures such that, for all $I \in \mathcal{T}$,

$$\mathbb{P}(\xi_1(I) = 0) = \mathbb{P}(\xi_2(I) = 0),$$

(2.4.22)

then

$$\xi_1^* \overset{d}{=} \xi_2^*.$$  \hspace{1cm} (2.4.23)

To see this, let

$$\mathcal{C} = \{ \{ \mu \in M_p(E) : \mu(I) = 0 \} ; I \in \mathcal{T} \}.$$  \hspace{1cm} (2.4.24)

The set $\mathcal{C}$ is easily seen to be a $\Pi$-system. Thus, since by assumption the laws, $\mathbb{P}_i$, of the point processes $\xi_i$ coincide on this $\Pi$-system, they coincide on the sigma-algebra generated by it. We must now check that $T^*$ is measurable as a map from $(M_p, \sigma(\mathcal{C}))$ to $(M_p, \mathbb{M}_p)$, which will hold, if for each $I$, the map $T^*_I : \mu \mapsto \mu^*(I)$ is measurable form $(M_p, \sigma(\mathcal{C})) \to \{0,1,2,\ldots\}$. Now introduce a family of finite coverings of (the relatively compact set) $I, A_{n,j}$, with $A_{n,j}$’s whose diameter is less than $1/n$. We will chose the family such that for each $j, A_{n+1,j} \subset A_{n,j}$, for some $i$. Then

$$T^*_I \mu = \mu^*(I) = \lim_{n \to \infty} \sum_{j=1}^{k_n} \mu(A_{n,j}) \wedge 1,$$

(2.4.25)

since eventually, no $A_{n,j}$ will contain more than one point of $\mu$. Now set $T^*_2 \mu = (\mu(A_{n,j}) \wedge 1)$. Clearly,

$$(T^*_2)^{-1}\{0\} = \{ \mu : \mu(A_{n,j}) = 0 \} \subset \sigma(\mathcal{C}),$$

(2.4.26)

and so $T^*_2$ is measurable as desired, and so is $T^*_1$, being a monotone limit of finite sums of measurable maps. But now

$$\mathbb{P}(\xi_1^* \in B) = \mathbb{P}(T^* \xi_1 \in B) = \mathbb{P}(\xi_1 \in (T^*)^{-1}(B)) = \mathbb{P}_1((T^*)^{-1}(B)).$$

(2.4.27)

But since $(T^*)^{-1}(B) \in \sigma(\mathcal{C})$, by hypothesis, $\mathbb{P}_1((T^*)^{-1}(B)) = \mathbb{P}_2((T^*)^{-1}(B))$, which is also equal to $\mathbb{P}(\xi_1^* \in B)$, which proves (2.4.22).

Now, as we have already mentioned, (2.4.18) implies tightness of the sequence $\xi_n$. Thus, for any subsequence $n'$, there exist a sub-sub-sequence, $n''$, such that $\xi_{n''}$ converges weakly to a limit, $\eta$. By Proposition 2.16 this is a point process. Let us assume for a moment that (a) $\eta$ is simple, and (b), for any relatively compact $A$,

$$\mathbb{P}(\xi^*(\partial A) = 0) \Rightarrow \mathbb{P}(\eta(\partial A) = 0).$$

(2.4.28)

Then, the map $\mu \mapsto \mu(I)$ is a.s. continuous with respect to $\eta$, and therefore, if $\xi_n^* \overset{w}{\to} \eta$, then

$$\mathbb{P}(\xi_n^*(I) = 0) \to \mathbb{P}(\eta(I) = 0).$$

(2.4.29)

But we assumed that

$$\mathbb{P}(\xi_n^*(I) = 0) \to \mathbb{P}(\xi(I) = 0),$$

(2.4.30)
so that, by the foregoing observation, and the fact that both \( \eta \) and \( \xi \) are simple, 
\[ \xi = \eta. \]

It remains to check simplicity of \( \eta \) and (2.4.28).

To verify the latter, we will show that for any compact set, \( K \),
\[ \mathbb{P}(\eta(K) = 0) \geq \mathbb{P}(\xi(K) = 0). \tag{2.4.31} \]

We use that for any such \( K \), there exist sequences of functions, \( f_j \in C^+_0(\mathbb{R}^d) \), and compact sets, \( K_j \), such that
\[ 1_K \leq f_j \leq 1_{K_j}, \tag{2.4.32} \]
and \( 1_{K_j} \downarrow 1_K \). Thus,
\[ \mathbb{P}(\eta(K) = 0) \geq \mathbb{P}(\eta(f_j) = 0) = \mathbb{P}(\eta(f_j) \leq 0). \tag{2.4.33} \]

But \( \xi_{n'}(f_j) \) converges to \( \eta(f_j) \), and so
\[ \mathbb{P}(\eta(f_j) \leq 0) \geq \limsup_{n'} \mathbb{P}(\xi_{n'}(f_j) \leq 0) \geq \mathbb{P}(\xi_{n'}(K_j) \leq 0). \tag{2.4.34} \]

Finally, we can approximate \( K_j \) by elements \( I_j \in \mathcal{I} \), such that \( K_j \subset I_j \subset K \), so that
\[ \mathbb{P}(\xi_{n'}(K_j) \leq 0) \geq \limsup_{n'} \mathbb{P}(\xi_{n'}(I_j) \leq 0) = \mathbb{P}(\xi(I) \leq 0), \tag{2.4.35} \]
so that (2.4.31) follows.

Finally, to show simplicity, we take \( I \in \mathcal{I} \) and show that the \( \eta \) has multiple points in \( I \) with zero probability. Now
\[ \mathbb{P}(\eta(I) > \eta^*(I)) = \mathbb{P}(\eta(I) - \eta^*(I) < 1/2) \leq 2(\mathbb{E}\eta(I) - \mathbb{E}\eta^*(I)). \tag{2.4.36} \]

\[ \square \]

The latter, however, is zero, due to the assumption of convergence of the intensity measures.

**Remark 2.25.** The main requirement in the theorem is the convergence of the so-called *avoidance function*, \( \mathbb{P}(\xi_{n'}(I) = 0) \). (2.4.18). The convergence of the mean (the intensity measure) provides tightness. It may be replaced by any other tightness criterion (see [30]). Note that, by Chebychev’s inequality, (2.4.18) implies tightness via Corollary 2.21. We will have to deal with situations where (2.4.18) fails, e.g. in branching Brownian motion.
2.5 Point processes of extremes

We are now ready to describe the structure of extremes of random sequences in terms of point processes. There are several aspects of these processes that we may want to capture:

(i) the distribution of the values largest values of the process; if \( u_n(x) \) is the scaling function such that \( \mathbb{P}(M_n \leq u_n(x)) \rightarrow G(x) \), it would be natural to look at the point process

\[
N_n \equiv \sum_{i=1}^{n} \delta_{u_n^{-1}(X_i)},
\]

(2.5.1)

As \( n \) tends to infinity, most of the points \( u_n(X_i) \) will disappear to minus infinity, but we may hope that as a point process, this object will converge.

(ii) the “spatial” structure of the large values: we may fix an extreme level, \( u_n \), and ask for the distribution of the values \( i \) for which \( X_i \) exceeds this level. Again, only a finite number of exceedances will be expected. To represent the exceedances as point process, it will be convenient to embed \( 1 \leq i \leq n \) in the unit interval \( (0,1] \), via the map \( i \rightarrow i/n \). This leads us to consider the point process of exceedances on \( (0,1] \),

\[
N_n \equiv \sum_{i=1}^{n} \delta_{i/n} 1_{X_i > u_n}.
\]

(2.5.2)

(iii) we may consider the two aspects together and consider the point process on \( \mathbb{R} \times (0,1] \),

\[
N_n \equiv \sum_{i=1}^{n} \delta_{u_n^{-1}(X_i),i/n}.
\]

(2.5.3)

In this chapter we consider only the case of iid sequences, although this is far too restrictive. We turn to the non-iid case later when we study Gaussian processes.

2.5.1 The point process of exceedances.

We begin with the simplest, object, the process \( N_n \) of exceedances of an extremal level \( u_n \).

**Theorem 2.26.** Let \( X_i \) be iid random variables with marginal distribution function \( F \). Let \( \tau > 0 \) and assume that there is \( u_n \equiv u_n(\tau) \) such that \( n(1 - F(u_n(\tau))) \rightarrow \tau \). then the point process

\[
\tilde{N}_n \equiv \sum_{i=\infty}^{\infty} \delta_{i/n} 1_{X_i > u_n(\tau)}
\]

(2.5.4)

converges weakly to the Poisson point process \( \tilde{N} \) on \( \mathbb{R} \) with intensity measure \( \tau dx \).

**Proof.** We will use Kallenberg’s theorem. First note that trivially,
2.5 Point processes of extremes

\[ \mathbb{E} N_n((c,d]) = \sum_{i=1}^{n} \mathbb{P}(X_i > u_n(\tau)) \mathbb{1}_{i/n \in (c,d]} \]  (2.5.5)

\[ = n(d-c)(1-F(u_n(\tau))) \to \tau(d-c), \]

so that the intensity measure converges to the desired one.

Next we need to show that

\[ \mathbb{P}(N_n(I) = 0) \to e^{-\tau|I|}, \]  (2.5.6)

for \( I \) any finite union of disjoint intervals. Then

\[ \mathbb{P}(N_n(I) = 0) = \mathbb{P}(\forall i/n \in I X_i \leq u_n) \to e^{-\tau|I|}, \]  (2.5.7)

from the basic result on convergence of the law of the maximum. \( \square \)

2.5.2 The point process of extreme values.

Let us now turn to an alternative point process, that of the values of the largest maxima. In the iid case we have already shown enough in Theorem 1.19 to get the following result. However, we will use the occasion to show how to use Laplace functionals.

**Theorem 2.27.** Let \( X_i \) be sequence of iid random variables with marginal distribution function \( F \). Assume that for all \( \tau \in \mathbb{R}^+ \), \( n(1-F(u_n(\tau))) \to \tau \), uniformly on compact intervals. Then the point process

\[ E_n \equiv \sum_{i=1}^{n} \delta_{u^{-1}_n(X_i)} \]  (2.5.8)

converges weakly to the Poisson Point process on \( \mathbb{R}^+ \) with intensity measure the Lebesgue measure.

**Proof.** We will show this by proving convergence of the Laplace functionals. For \( \phi : \mathbb{R} \to \mathbb{R}^+ \) a continuous, non-negative function of compact support, we set

\[ \Psi_n(\phi) \equiv \mathbb{E} \left[ \exp \left( - \int \phi(x)E_n(dx) \right) \right] = \mathbb{E} \left[ \exp \left( - \sum_{i=1}^{n} \phi \left( u^{-1}_n(X_i) \right) \right) \right]. \]  (2.5.9)

The computations in the iid case are very simple. By independence,

\[ \mathbb{E} \left[ \exp \left( - \sum_{i=1}^{n} \phi \left( u^{-1}_n(X_i) \right) \right) \right] = \prod_{i=1}^{n} \mathbb{E} \left[ \exp \left( - \phi \left( u^{-1}_n(X) \right) \right) \right] \]

\[ = (\mathbb{E} \left[ \exp \left( - \phi \left( u^{-1}_n(X) \right) \right) \right])^n. \]  (2.5.10)
We know that the probability for $u_n(X_i)$ to be in the support of $\phi$ is small of order $1/n$. Thus, if we write
\[
\mathbb{E} \left[ e^{-\phi(u_n^{-1}(X_i))} \right] = 1 + \mathbb{E} \left[ e^{-\phi(u_n^{-1}(X_i))} - 1 \right],
\]
the second term will be of order $1/n$. Therefore,
\[
\left( \mathbb{E} \left[ e^{-\phi(u_n^{-1}(X_i))} \right] \right)^n \sim \exp \left( n \mathbb{E} \left[ e^{-\phi(u_n^{-1}(X_i))} - 1 \right] \right). \tag{2.5.12}
\]
Finally,
\[
\lim_{n \to \infty} n \mathbb{E} \left[ e^{-\phi(u_n^{-1}(X_i))} - 1 \right] = \int \left( e^{-\phi(\tau)} - 1 \right) d\tau. \tag{2.5.13}
\]
To show the latter, note first that, (assuming first that $\phi$ is differentiable and using integration by parts,)
\[
n \mathbb{E} \left[ e^{-\phi(u_n^{-1}(X_i))} - 1 \right] = \int^\infty_0 n \mathbb{P}(u_n^{-1}(X_1) \in d\tau) \left( e^{-\phi(\tau)} - 1 \right) \tag{2.5.14}
\]
\[
= \int^\infty_0 \frac{d}{d\tau} \left( e^{-\phi(\tau)} - 1 \right) \mathbb{P}(u_n^{-1}(X_1) > \tau). \]

Since $\phi$ has compact support and the integrand converges uniformly on compact sets, the assertion (2.5.13) follows. A standard approximation argument shows that the same holds for all $\phi \in C^+_0$. From (2.5.13) we get that the Laplace functional converges to that of the PPP$(d\tau)$, which proves the theorem.

### 2.5.3 Complete Poisson convergence.

We now come to the final goal of this section, the characterisation of the space-value process of extremes as a two-dimensional point process. We consider again $u_n(\tau)$ such that $n(1 - F(u_n(\tau))) \to \tau$. Then we define
\[
\mathcal{N}_n \equiv \sum_{i=1}^\infty \delta_{(i/n,u_n^{-1}(X_i))}, \tag{2.5.15}
\]
as a point process on $\mathbb{R}^2$ (or more precisely, on $\mathbb{R}_+ \times \mathbb{R}_+$).

**Theorem 2.28.** Let $u_n(\tau)$ be as above. Then the point process $\mathcal{N}_n$ converges to the Poisson point process, $\mathcal{N}$, on $\mathbb{R}_+^2$ with intensity measure given by the Lebesgue measure.

**Proof.** The easiest way in the iid case to prove this theorem is to use Laplace functionals just as in the proof of Theorem 2.27. For $\phi \in C^+_0(\mathbb{R}_+^2)$. Clearly
\[ \Psi_{N_n}(\phi) = \prod_{i=1}^{\infty} (1 + E \left[ \exp \left( -\phi \left( u_n^{-1}(X_i), i/n \right) \right) - 1 \right]) \quad (2.5.16) \]

\[ \sim \exp \left( n \sum_{i=1}^{\infty} nE \left[ \exp \left( -\phi \left( u_n^{-1}(X_i), i/n \right) \right) - 1 \right] \right). \]

Note that since \( \phi \) has compact support, there are only finitely many non-zero terms in the sum. As before,

\[ nE \left[ \exp \left( -\phi \left( u_n^{-1}(X_i), i/n \right) \right) - 1 \right] = \int_0^{\infty} \left( e^{-\phi(\tau, i/n)} - 1 \right) d\tau \quad (2.5.17) \]

\[ \int_0^{\infty} \frac{d}{d\tau} \left( e^{-\phi(\tau, i/n)} - 1 \right) \left( nP \left( u_n^{-1}(X_i) > \tau \right) - \tau \right) d\tau. \]

The first term is what we want, while again due to the fact that \( \phi \) has compact support (and for the moment has bounded derivatives), say \( I \times J \) the second term is in absolute value smaller than

\[ \mathbb{1}_{i/n \in J} \int_I \left( nP \left( u_n^{-1}(X_i) \leq \tau \right) - \tau \right) d\tau. \quad (2.5.18) \]

Since the term in the bracket tends to zero, inserting this into the sum over \( i \) still gives a vanishing contribution, while

\[ n^{-1} \sum_{i=1}^{\infty} \int_0^{\infty} \left( e^{-\phi(\tau, i/n)} - 1 \right) d\tau \rightarrow \int_0^{\infty} \int_0^{\infty} \left( e^{-\phi(\tau, z)} - 1 \right) d\tau dz. \quad (2.5.19) \]

From here the claimed result is obvious. \( \square \)
Chapter 3
Normal sequences

The assumption that the underlying random process are iid is, of course rather restrictive and unrealistic in real applications. A lot of work has been done to extend extreme value theory to other processes. It turns out that the results of the iid model are fairly robust, and survive essentially unchanged under relatively weak mixing assumptions. In this course we are mainly interested in the case of Gaussian processes and for this reason we present some of the main classical results for this case only in the present chapter. More precisely, we consider the case of stationary Gaussian sequences. We will introduce a very powerful tool, Gaussian comparison, that makes the study of extremes in the Gaussian case more amenable.

In the stationary case, a normalised Gaussian sequence, \((X_i, i \in \mathbb{Z})\), is characterised by

\[
\begin{align*}
\mathbb{E}X_i &= 0, \\
\mathbb{E}X_i^2 &= 1, \\
\mathbb{E}X_iX_j &= r_{i-j},
\end{align*}
\]

where \(r_k = r_{|k|}\). The sequence must of course be such that the infinite dimensional matrix with entries \(c_{ij} = r_{i-j}\) is positive definite.

Our main target here is to show that under the so-called Berman condition \((3.0.2)\),

\[
r_q \ln n \downarrow 0,
\]

the extremes of a stationary normal sequences behave like those of the corresponding iid normal sequence. We shall see that the logarithmic decay of correlations is indeed a boundary for irrelevance of correlation.
3.1 Normal comparison

In the context of Gaussian random variables, a recurrent idea is to compare one Gaussian process to another, simpler one. The simplest one to compare with are, of course, iid variables, but the concept goes much farther.

Let us consider a family of Gaussian random variables, \( \xi_1, \ldots, \xi_n \), normalised to have mean zero and variance one (we refer to such Gaussian random variables as centred normal random variables), and let \( \Lambda^1 \) denote their covariance matrix. Let similarly \( \eta_1, \ldots, \eta_n \) be centred normal random variables with covariance matrix \( \Lambda^0 \).

Generally speaking, one is interested in comparing functions of these two processes that typically will be of the form

\[
E F(X_1, \ldots, X_n),
\]

(3.1.1)

where \( F : \mathbb{R}^n \to \mathbb{R} \). For us the most common case would be

\[
F(X_1, \ldots, X_n) = 1_{X_1 \leq x_1, \ldots, X_n \leq x_n}.
\]

(3.1.2)

An extraordinary efficient tool to compare such processes turns out to be interpolation. Given \( \xi \) and \( \eta \), we define \( X^h_1, \ldots, X^h_n \), for \( h \in [0, 1] \), by

\[
X^h = \sqrt{h} \xi + \sqrt{1-h} \eta.
\]

(3.1.3)

Then \( X^h \) is normal and has covariance

\[
\Lambda^h = h\Lambda^1 + (1-h)\Lambda^0,
\]

(3.1.4)

The following Gaussian comparison lemma is a fundamental tool in the study of Gaussian processes.

**Lemma 3.1.** Let \( \eta, \xi, X^h \) be as above. Let \( F : \mathbb{R}^n \to \mathbb{R} \) be differentiable and of moderate growth. Set \( f(h) \equiv E F(X^h_1, \ldots, X^h_n) \). Then

\[
f(1) - f(0) = \frac{1}{2} \int_0^1 dh \sum_{i \neq j} (\Lambda^1_{ij} - \Lambda^0_{ij}) E \left( \frac{\partial^2 F}{\partial x_i \partial x_j}(X^h_1, \ldots, X^h_n) \right).
\]

(3.1.5)

**Proof.** Trivially,

\[
f(1) - f(0) = \int_0^1 dh \frac{d}{dh} f(h),
\]

(3.1.6)

and

\[
\frac{d}{dh} f(h) = \frac{1}{2} \sum_{i=1}^n E \left( \frac{\partial F}{\partial x_i} \left( h^{-1/2} \xi_i (1 - h)^{-1/2} \eta_i \right) \right),
\]

(3.1.7)

where of course \( \frac{\partial F}{\partial x_i} \) is evaluated at \( X^h \). To continue we use a remarkable formula for Gaussian processes, known as the Gaussian integration by parts formula.
Lemma 3.2. Let \( X_i, i \in \{1, \ldots, n\} \) be a multivariate Gaussian process, and let \( g: \mathbb{R}^n \to \mathbb{R} \) be a differentiable function of at most polynomial growth. Then

\[
\mathbb{E}g(X)X_i = \sum_{j=1}^{n} \mathbb{E}(X_iX_j)\mathbb{E} \frac{\partial}{\partial x_j} g(X). \tag{3.1.8}
\]

Proof. We first consider the scalar case, i.e.

Lemma 3.3. Let \( X \) be a centred Gaussian random variable, and let \( g: \mathbb{R} \to \mathbb{R} \) be a differentiable function of at most polynomial growth. Then

\[
\mathbb{E}g(X)X = \mathbb{E}(X^2)\mathbb{E}g'(X). \tag{3.1.9}
\]

Proof. Let \( \sigma = \mathbb{E}X^2 \). Then

\[
\mathbb{E}Xg(X) = \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} g(x)e^{-\frac{x^2}{2\sigma^2}} dx
= \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} g(x) \frac{d}{dx} \left( -\sigma^2 e^{-\frac{x^2}{2\sigma^2}} \right) dx
= \sigma^2 \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} \frac{d}{dx} \left( g(x)e^{-\frac{x^2}{2\sigma^2}} \right) dx = \mathbb{E}X^2 \mathbb{E}g'(x), \tag{3.1.10}
\]

where we used elementary integration by parts and the assumption that \( g(x)e^{-\frac{x^2}{2\sigma^2}} \downarrow 0 \), if \( x \to \pm \infty \). \( \square \)

To prove the multivariate case, we use a trick from [71]: We let \( X_i, i = 1, \ldots, n \) be a centred Gaussian vector, and let \( X \) be a centred Gaussian random variable. Set \( X'_i = X_i - X \mathbb{E}XX_i \mathbb{E}X^2 \). Then

\[
\mathbb{E}X'_iX = \mathbb{E}X_iX - \mathbb{E}X_iX = 0, \tag{3.1.11}
\]

and so \( X \) is independent of the vector \( X'_i \). Now compute

\[
\mathbb{E}XF(X_1, \ldots, X_n) = \mathbb{E}XF \left( X'_1 + X \frac{\mathbb{E}X_iX}{\mathbb{E}X^2}, \ldots, X'_n + X \frac{\mathbb{E}X_nX}{\mathbb{E}X^2} \right). \tag{3.1.12}
\]

Using in this expression Lemma 3.3 for the random variable \( X \) alone, we obtain

\[
\mathbb{E}XF(X_1, \ldots, X_n) = \mathbb{E}X^2\mathbb{E}F' \left( X'_1 + X \frac{\mathbb{E}X_iX}{\mathbb{E}X^2}, \ldots, X'_n + X \frac{\mathbb{E}X_nX}{\mathbb{E}X^2} \right)
= \mathbb{E}X^2 \sum_{i=1}^{n} \mathbb{E}X_iX \frac{\partial F}{\partial x_i} \left( X'_1 + X \frac{\mathbb{E}X_iX}{\mathbb{E}X^2}, \ldots, X'_n + X \frac{\mathbb{E}X_nX}{\mathbb{E}X^2} \right)
= \sum_{i=1}^{n} \mathbb{E}(X_iX) \mathbb{E} \frac{\partial F}{\partial x_i} (X_1, \ldots, X_n),
\]

which proves the lemma. \( \square \)
Applying Lemma (3.2) in (3.1.7) yields

\[ \frac{d}{dh} f(h) = \frac{1}{2} \sum_{i \neq j} \mathbb{E} \frac{\partial^2 F}{\partial x_j \partial x_i} \mathbb{E} (\xi_i \xi_j - \eta_i \eta_j) \]  

\[ = \frac{1}{2} \sum_{i \neq j} (A_{ji}^1 - A_{ji}^0) \mathbb{E} \frac{\partial^2 F}{\partial x_j \partial x_i} \left( X_{1j}^h, \ldots, X_{nj}^h \right), \]

which is the desired formula. \( \square \)

The general comparison lemma can be put to various good uses. The first is a monotonicity result that is sometimes known as Kahane’s theorem [49].

**Theorem 3.4.** Let \( \xi \) and \( \eta \) be two independent \( n \)-dimensional Gaussian vectors. Let \( D_1 \) and \( D_2 \) be subsets of \( \{1, \ldots, n\} \times \{1, \ldots, n\} \). Assume that

\[ \mathbb{E} \xi_i \xi_j \geq \mathbb{E} \eta_i \eta_j, \quad \text{if} \quad (i, j) \in D_1, \]
\[ \mathbb{E} \xi_i \xi_j \leq \mathbb{E} \eta_i \eta_j, \quad \text{if} \quad (i, j) \in D_2, \]
\[ \mathbb{E} \xi_i \xi_i = \mathbb{E} \eta_i \eta_i, \quad \text{if} \quad (i, j) \notin D_1 \cup D_2. \]  

(3.1.14)

Let \( F \) be a function on \( \mathbb{R}^n \) of moderate growth, such that its second derivatives satisfy

\[ \frac{\partial^2}{\partial x_i \partial x_j} F(x) \geq 0, \quad \text{if} \quad (i, j) \in D_1, \]
\[ \frac{\partial^2}{\partial x_i \partial x_j} F(x) \leq 0, \quad \text{if} \quad (i, j) \in D_2. \]  

(3.1.15)

Then

\[ \mathbb{E} f(\xi) \leq \mathbb{E} f(\eta). \]  

(3.1.16)

**Proof.** The proof of the theorem can be trivially read off the preceding lemma by inserting the hypotheses into the right-hand side of (3.1.5). \( \square \)

We will need two extensions of these results for functions that are not differentiable. The first is known as Slepian’s lemma [69].

**Lemma 3.5.** Let \( \xi \) and \( \eta \) be two independent \( n \)-dimensional Gaussian vectors. Assume that

\[ \mathbb{E} \xi_i \xi_j \geq \mathbb{E} \eta_i \eta_j, \quad \text{for all} \quad i \neq j \]
\[ \mathbb{E} \xi_i \xi_i = \mathbb{E} \eta_i \eta_i, \quad \text{for all} \quad i. \]  

(3.1.17)

Then

\[ \mathbb{E} \max_{i=1}^n (\xi_i) \leq \mathbb{E} \max_{i=1}^n (\eta_i). \]  

(3.1.18)
Proof. Let

\[ F_\beta(x_1, \ldots, x_n) \equiv \beta^{-1} \ln \sum_{i=1}^n e^{\beta x_i}. \] (3.1.19)

A simple computation shows that, for \( i \neq j \),

\[ \frac{\partial^2 F}{\partial x_i \partial x_j} = -\beta e^{\beta (x_i + x_j)} \left( \sum_{k=1}^n e^{\beta x_k} \right)^2, < 0, \] (3.1.20)

and so the theorem implies that, for all \( \beta > 0 \),

\[ \mathbb{E} F_\beta(\xi) \leq \mathbb{E} F_\beta(\eta). \] (3.1.21)

On the other hand,

\[ \lim_{\beta \to \infty} F_\beta(x_1, \ldots, x_n) = \max_{i=1}^n x_i. \] (3.1.22)

Thus

\[ \lim_{\beta \to \infty} \mathbb{E} F_\beta(\xi_1, \ldots, \xi_n) \leq \lim_{\beta \to \infty} \mathbb{E} F_\beta(\eta_1, \ldots, \eta_n), \] (3.1.23)

and hence (3.1.18) holds.

As a second application, we want to study \( \mathbb{P}(X_1 \leq u_1, \ldots, X_n \leq u_n) \). This corresponds to choosing \( F(X_1, \ldots, X_n) = \mathbb{1}_{X_1 \leq u_1, \ldots, X_n \leq u_n} \).

Lemma 3.6. Let \( \xi, \eta \) be as above. Set \( \rho_{ij} \equiv \max(A_{ij}, A_{ij}') \), and denote by \( x_+ \equiv \max(x, 0) \). Then

\[ \mathbb{P}(\xi_1 \leq u_1, \ldots, \xi_n \leq u_n) - \mathbb{P}(\eta_1 \leq u_1, \ldots, \eta_n \leq u_n) \]

\[ \leq \frac{1}{2\pi} \sum_{1 \leq i < j \leq n} \frac{A_{ij} + A_{ij}'}{\sqrt{1 - \rho_{ij}^2}} \exp \left( -\frac{u_i^2 + u_j^2}{2(1 + \rho_{ij})} \right). \] (3.1.24)

Proof. Although the indicator function is not differentiable, we will proceed as if it was, setting

\[ \frac{d}{dx} \mathbb{1}_{x \leq u} = \delta(x - u), \] (3.1.25)

where \( \delta \) denotes the Dirac delta function, i.e. \( \int f(x) \delta(x - u)dx \equiv f(u) \). This can be justified e.g. by using smooth approximants of the indicator function and passing to the limit at the end (e.g. replace \( \mathbb{1}_{x \leq u} \) by \( (2\pi\sigma^2)^{1/2} \int_{-\infty}^x \exp \left( -\frac{(z-u)^2}{2\sigma^2} \right) dz \)), do all the computations, and pass to the limit \( \sigma \downarrow 0 \) at the end. With this convention, we have that, for \( i \neq j \),

\[ \mathbb{E} \left( \frac{\partial^2 F}{\partial x_i \partial x_j}(X_1^h, \ldots, X_n^h) \right) = \mathbb{E} \prod_{k \neq i, j} \mathbb{1}_{X_k \leq u_k} \delta(X_i^h - u_i) \delta(X_j^h - u_j). \] (3.1.26)

\[ \leq \mathbb{E} \delta(X_i^h - u_i) \delta(X_j^h - u_j) = \phi_i(u_i, u_j), \]
where $\phi_h$ denotes the density of the bivariate normal distribution with covariance $\Lambda^h_{ij}$, i.e.

$$
\phi_h(u_i, u_j) = \frac{1}{2\pi\sqrt{1-(\Lambda^h_{ij})^2}} \exp \left( -\frac{u_i^2 + u_j^2 - 2\Lambda^h_{ij}u_iu_j}{2(1-(\Lambda^h_{ij})^2)} \right),\tag{3.1.27}
$$

Now

$$
u_i^2 + u_j^2 - 2\Lambda^h_{ij}u_iu_j \geq \frac{(u_i^2 + u_j^2)(1 - \Lambda^h_{ij}) + \Lambda^h_{ij}(u_i - u_j)^2}{2(1-(\Lambda^h_{ij})^2)} \geq \frac{(u_i^2 + u_j^2)}{2(1+|\Lambda^h_{ij}|)} \geq \frac{(u_i^2 + u_j^2)}{2(1+\rho_{ij})},\tag{3.1.28}
$$

where $\rho_{ij} = \max(\Lambda^0_{ij}, \Lambda^1_{ij})$. (To prove the first inequality, note that this is trivial if $\Lambda^h_{ij} \geq 0$. If $\Lambda^h_{ij} < 0$, the result follows after some simple algebra). Inserting this into (3.1.26) gives

$$E\left( \frac{\partial^2 F}{\partial x_i \partial x_j}(X^h_1, \ldots, X^h_n) \right) \leq \frac{1}{2\pi \sqrt{1-\rho_{ij}^2}} \exp \left( -\frac{(u_i^2 + u_j^2)}{2(1+\rho_{ij})} \right),\tag{3.1.29}
$$

from which (3.1.24) follows immediately. $\Box$

**Remark 3.7.** It is often convenient to replace the assertion of Lemma 3.6 by

$$|\mathbb{P}(\xi_1 \leq u_1, \ldots, \xi_n \leq u_n) - \mathbb{P}(\eta_1 \leq u_1, \ldots, \eta_n \leq u_n)| \leq \frac{1}{2\pi} \sum_{1 \leq i < j \leq n} \frac{|\Lambda^1_{ij} - \Lambda^0_{ij}|}{\sqrt{1-\rho^2_{ij}}} \exp \left( -\frac{u_i^2 + u_j^2}{2(1+\rho_{ij})} \right).\tag{3.1.30}
$$

A simple, but useful corollary is the specialisation of this lemma to the case when $\eta_i$ are independent random variables.

**Corollary 3.8.** Let $\xi_i$ be centred normal variables with covariance matrix $\Lambda$, and let $\eta_i$ be iid centred normal variables. Then

$$\mathbb{P}(\xi_1 \leq u_1, \ldots, \xi_n \leq u_n) - \mathbb{P}(\eta_1 \leq u_1, \ldots, \eta_n \leq u_n) \leq \frac{1}{2\pi} \sum_{1 \leq i < j \leq n} \frac{(\Lambda_{ij})_+}{\sqrt{1-\Lambda^2_{ij}}} \exp \left( -\frac{u_i^2 + u_j^2}{2(1+|\Lambda_{ij}|)} \right).\tag{3.1.31}
$$

In particular, if $|\Lambda_{ij}| < \delta \leq 1$. 
\[ | \mathbb{P}(\xi_1 \leq u, \ldots, \xi_n \leq u) - [\Phi(u)]^n | \leq \frac{1}{2\pi} \sum_{1 \leq i < j \leq n} \frac{|A_{ij}|}{\sqrt{1-\delta^2}} \exp \left( -\frac{u^2}{1+|A_{ij}|} \right). \] 

**Proof.** The proof of the corollary is straightforward and left to the reader. \( \square \)

Another simple corollary is a version of Slepian’s lemma:

**Corollary 3.9.** Let \( \xi, \eta \) be as above. Assume that \( A_{ij}^0 \leq A_{ij}^1 \), for all \( i \neq j \). Then

\[ \mathbb{P}\left( \max_{i=1}^n \xi \leq u \right) - \mathbb{P}\left( \max_{i=1}^n \eta \leq u \right) \leq 0. \] 

**Proof.** The proof of the corollary is again obvious, since under our assumption, \( (A_{ij}^0 - A_{ij}^1)_+ = 0 \). \( \square \)

### 3.2 Applications to extremes

The comparison results of Section 3.1 can readily be used to give criterion under which the extremes of correlated Gaussian sequences are distributed as in the independent case.

**Lemma 3.10.** Let \( \xi_i, i \in \mathbb{Z} \) be a stationary normal sequence with covariance \( r_n \). Assume that \( \sup_{n \geq 1} r_n \leq \delta < 1 \). Let \( u_n \) be such that

\[ \lim_{n \to \infty} n \sum_{i=1}^n |r_i| e^{-\frac{u_n^2}{1+|r_i|}} = 0. \] 

Then

\[ n (1 - \Phi(u_n)) \to \tau \iff \mathbb{P}(M_n \leq u_n) \to e^{-\tau}. \] 

**Proof.** Using Corollary 3.8, we see that (3.2.1) implies that

\[ \mathbb{P}(M_n \leq u_n) - \Phi(u_n)^n \downarrow 0. \] 

Since

\[ n (1 - \Phi(u_n)) \to \tau \iff \Phi(u_n)^n \to e^{-\tau}, \] 

the assertion of the lemma follows. \( \square \)

Since the condition \( n (1 - \Phi(u_n)) \to \tau \) determines \( u_n \) (if \( 0 < \tau < \infty \)), on can easily derive a criteria for (3.2.1) to hold.

**Lemma 3.11.** Assume that \( r_n \ln n \downarrow 0 \), and that \( u_n \) is such that \( n (1 - \Phi(u_n)) \to \tau \), \( 0 < \tau < \infty \). Then (3.2.1) holds.
Proof. We know that, if \( n(1 - \Phi(u_n)) \sim \tau \),

\[
\exp \left( - \frac{1}{2} u_n^2 \right) \sim Ku_n n^{-1}.
\]

(3.2.5)

and \( u_n \sim \sqrt{2 \ln n} \). Thus

\[
n|r_i|e^{-\frac{u_n^2}{2|\ln n|}} = n|r_i|e^{-u_n^2}e^{\frac{u_n^2}{2|\ln n|}}.
\]

(3.2.6)

Let \( \alpha > 0 \), and \( i \geq n^\alpha \). Then

\[
n|r_i|e^{-u_n^2} \sim 2n^{-1}|r_i|\ln n,
\]

(3.2.7)

and

\[
\frac{u_n^2|r_i|}{1 + |r_i|} \leq 2|r_i|\ln n.
\]

(3.2.8)

But then

\[
|r_i|\ln n = |r_i|\ln \frac{\ln n}{\ln i} \leq |r_i|\ln \frac{\ln n}{\ln n^\alpha} = \alpha^{-1}|r_i|\ln i,
\]

(3.2.9)

which tends to zero, since \( r_n\ln n \to 0 \), if \( i \to \infty \). Thus

\[
\sum_{i \geq n^\alpha} n|r_i|e^{-\frac{u_n^2}{|\ln n|}} \leq 2\alpha^{-1} \sup_{i \geq n^\alpha} |r_i|\ln i \exp \left( 2\alpha^{-1}|r_i|\ln i \right) \downarrow 0,
\]

(3.2.10)

as \( n \to \infty \). On the other hand, since there exists \( \delta > 0 \), such that \( 1 - r_i \geq \delta \),

\[
\sum_{i \leq n^\alpha} n|r_i|e^{-\frac{u_n^2}{|\ln n|}} \leq n^{1 + \alpha} n^{-2/(2 - \delta)} (2\ln n)^2,
\]

(3.2.11)

which tends to zero as well, provided we chose \( \alpha \) such that \( 1 + \alpha < 2/(2 - \delta) \), i.e. \( \alpha < \delta/(2 + \delta) \). This proves the lemma. \( \Box \)

The following theorem summarises the results on the stationary Gaussian case.

**Theorem 3.12.** Let \( X_i, i \in \mathbb{N} \) be a stationary centred normal series with covariance \( r_n \), such that \( r_n\ln n \to 0 \). Then

(i) For \( 0 \leq \tau \leq \infty \),

\[
n(1 - \Phi(u_n)) \to \tau \iff \Phi(u_n)^n \to e^{-\tau},
\]

(3.2.12)

(ii) \( u_n(x) \) can be chosen as in the iid normal case, and

(iii) The extremal process is Poisson, more precisely,

\[
\sum_{i=1}^n \delta_{u^{-1}_n(x_i)} \to \text{PPP}(e^{-x}dx),
\]

(3.2.13)
3.2 Applications to extremes

where PPP($\mu$) denotes the Poisson point process on $\mathbb{R}$ with intensity measure $\mu$.

**Proof.** Points (i) and (ii) have already been show. For the proof of the convergence of extremal process, we show the convergence of the Laplace functional. Let $\phi : \mathbb{R} \to \mathbb{R}_+$ have compact support (say in $[a, b]$), and assume without loss that $\phi$ has bounded derivatives. Then, by Lemma 3.1

$$E\left[e^{-\sum_{i=1}^{n} \phi(u_n^{-1}(X_i))}\right] = \prod_{i=1}^{n} E\left[e^{-\phi(u_n^{-1}(X_i))}\right],$$

(3.2.14)

By assumption on the function $\phi$, the right-hand side of (3.2.14) is bounded in absolute value by

$$C \sum_{i \neq j} A_{ij} \mathbb{P}\left(X_i^h > u_N(a), X_j^h > u(a)\right) 2 \ln n,$$

(3.2.15)

for some constant $C$. Using the formula for the joint density of the variables $X_i^h, X_j^h$ and the bound (3.1.28), we see that

$$\mathbb{P}\left(X_i^h > u_N(a), X_j^h > u(a)\right) \leq \frac{1}{2\pi \sqrt{(1 - A_{ij})}} \int_{u_n(a)}^{\infty} dx \int_{u_n(a)}^{\infty} dy \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} e^{-\frac{x^2 + y^2}{2(1 - A_{ij})}},$$

$$\leq \frac{(1 + A_{ij})}{2\pi u_n(a)^2 \sqrt{(1 - A_{ij})}} e^{-\frac{\pi u_n(a)^2}{(1 - A_{ij})}}.$$  

(3.2.16)

Using that $u_n^2(a) \sim \ln n$ and inserting this bound into (3.2.15), we see that (3.2.14) tends to zero as $n \uparrow \infty$, whenever the assumption of Lemma 3.10 is satisfied. This proves the theorem.  \[ \Box \]
Chapter 4
Spin glasses

The motivation for the results I present in these lectures comes from spin glass theory. There is really not enough time to go into this at any depth and I will only sketch a few key concepts. Those interested in more should read a book on the subject, e.g. [15], [71] or [64], or the less technical introduction by Newman and Stein [62].

4.1 Setting and examples

Spin glasses are spin systems with competing, random interactions. To remain close to the main frame of these lectures, let us remain in the realm of mean-field spin glasses of the type proposed by Sherrington and Kirkpatrick.

Here we have a state space, \( S_N \equiv \{-1, 1\}^N \). On this space we define a Gaussian process \( H_N : S_N \rightarrow \mathbb{R} \), characterised by its mean \( \mathbb{E} H_N(\sigma) = 0 \), for all \( \sigma \in S_N \), and covariance

\[
\mathbb{E} H_N(\sigma) H_N(\sigma') = N g(\sigma, \sigma'), \quad \sigma, \sigma' \in S_N,
\]

where \( g \) is a positive definite quadratic form. The random functions \( H_N \) are called Hamiltonians. In physics, they represent the energy of a configuration of spins, \( \sigma \).

In the examples that are most popular, \( g \) is assumed to depend only on some distance between \( \sigma \) and \( \sigma' \). There are two popular distances:

- The Hamming distance,

\[
d_{\text{ham}}(\sigma, \sigma') = \sum_{i=1}^N \mathbb{1}_{\sigma_i \neq \sigma'_i} = \frac{1}{2} \left( N - \sum_{i=1}^N \sigma_i \sigma'_i \right). \tag{4.1.2}
\]

In this case one choses

\[
g(\sigma, \sigma') = p \left( 1 - 2d_{\text{ham}}(\sigma, \sigma')/N \right), \tag{4.1.3}
\]
with \( p \) a polynomial with non-negative coefficients. The most popular case is \( p(x) = x^2 \). One denotes the quantity

\[
1 - 2d_{\text{ham}}(\sigma, \sigma')/N \equiv R_N(\sigma, \sigma')
\]

(4.1.4)
as the overlap between \( \sigma \) and \( \sigma' \). The resulting class of models are called Sherrington-Kirkpatrick models.

- A second popular distance is the lexicographic distance

\[
d_{\text{lex}}(\sigma, \sigma') = N + 1 - \min(i : \sigma_i \neq \sigma'_i),
\]

(4.1.5)

Note the \( d_{\text{lex}} \) is an ultrametric. In this case,

\[
g(\sigma, \sigma') = A(1 - N^{-1}d_{\text{lex}}(\sigma, \sigma')).
\]

(4.1.6)

\( A : [0, 1] \to [0, 1] \) can here be chosen as any non-decreasing function. It will be convenient to denote \( q_N(\sigma, \sigma') \equiv N^{-1}(\min(i : \sigma_i \neq \sigma'_i) - 1) \) and call it the ultrametric overlap. The models obtained in this way were introduced by Derrida and Gardner \([38, 39]\) and are called generalised random energy models (GREM).

The GREMs have a more explicit representation in term of iid normal random variables. Let us fix a number, \( k \in \mathbb{N} \), of levels. Then introduce \( k \) natural numbers, \( 0 < \ell_1 < \ell_2 < \cdots < \ell_n = N \), and \( k \) real numbers \( a_1, \ldots, a_k \), such that \( a_1^2 + a_2^2 + \ldots + a_n^2 = 1 \). Finally, let \( \{X_{\sigma}^m\}^{1 \leq m \leq k}_{\sigma \in \{-1, 1\}^{\ell_1+\cdots+\ell_m}} \) be iid centred normal random variables with variance \( N \). The Hamiltonian of a \( k \)-level GREM is then given as a random function \( H_N : \{-1, 1\}^N \to \mathbb{R} \), defined as

\[
H_N(\sigma) \equiv \sum_{m=1}^k a_m X_{\sigma_1 \ldots \sigma_m}^m,
\]

(4.1.7)

where we write \( \sigma \equiv \sigma_1 \sigma_2 \ldots \sigma_k \) and \( \sigma_k \in \{-1, 1\}^{\ell_m} \). It is straightforward to see that \( H_N(\sigma) \) is Gaussian with variance \( N \), and that

\[
\mathbb{E}(H_N(\sigma)H_N(\tau)) = N \sum_{m=1}^k a_m^2 \mathbb{1}_{\ell_1 + \cdots + \ell_m}.
\]

(4.1.8)

Thus, if we define the function

\[
\Sigma(z) = N \sum_{m=1}^k a_m^2 \mathbb{1}_{\ell_1 + \cdots + \ell_m \leq z}, \quad z \in [0, N],
\]

(4.1.9)

, so that (4.1.8) can be written as

\[
\mathbb{E}(H_N(\sigma)H_N(\tau)) = \Sigma^2(\min(i : \sigma_i \neq \tau_i) - 1).
\]

(4.1.10)

This is the form of the correlation given previously, with \( \Sigma^2(z) = NA(N^{-1}z) \). Clearly the representation (4.1.7) is very useful for doing explicit computations.
The main objects one studies in statistical mechanics are the Gibbs measures associated to these Hamiltonians. They are probability measures on \( S_N \) that assign to \( \sigma \in S_N \) the probability

\[
\mu_{\beta,N}(\sigma) \equiv \frac{e^{-\beta H_N(\sigma)}}{Z_{\beta,N}},
\]

where the normalising factor \( Z_{\beta,N} \) is called the partition function. The parameter \( \beta \) is called the inverse temperature. Note that these measures are random variables on the underlying probability space \( (\Omega, \mathcal{F}, P) \) on which the Gaussian processes \( H_N \) are defined.

The objective of statistical mechanics is to understand the geometry of these measures for very large \( N \). This problem is usually interesting in the case when \( \beta \) is large, where the Gibbs measures will feel the geometry of the random process \( H_N \). In particular, one should expect that for large enough \( \beta \) (and finite \( N \)), \( \mu_{\beta,N} \) will assign mass only to configurations of almost minimal energy.

One way to acquire information on the Gibbs measures is thus to first analyse the structure of the minima of \( H_N \).

### 4.2 The REM

In order to illustrate what can go on in spin glasses, Derrida had the bright idea to introduce a simple toy model, the random energy model (REM). In the REM, the values \( H_N(\sigma) \) of the Hamiltonian are simply independent Gaussian random variables with mean zero and variance \( N \). One might think that this should be too simple, but, remarkably, quite a lot can be learned from this example.

Understanding the structure of the ground states in this model turns into the classical problem of extreme value theory for iid random variables, which is of course extremely well understood (there are numerous textbooks of which I like [57] and [65] best). We will set \( H_N(\sigma) \equiv -X_\sigma \).

The first question is for the value of the maximum of the \( X_\sigma, \sigma \in \mathcal{S}_N \).

#### 4.2.1 Rough estimates, the second moment method

To get a first feeling we ask for the right order to the maximum. More precisely, we ask for the right choice of functions \( u_N : \mathbb{R} \to \mathbb{R} \), such that the maximum is of that order \( u_N(x) \) with positive, \( x \)-dependent probability. To do so, introduce the counting function

\[
M_N(x) \equiv \sum_{\sigma \in \mathcal{S}_N} 1_{X_\sigma > u_N(x)},
\]

which is the number of \( \sigma \)'s such that \( X_\sigma \) exceeds \( u_N(x) \). Then
This may be called a first moment estimate. So a good choice of $u_N(x)$ would be such that $2^N P(X_\sigma > u_N(x)) = O(1)$. For later use, we even choose

$$2^N P(X_\sigma > u_N(x)) \sim e^{-x}. \quad (4.2.3)$$

It is easy to see that this can be achieved with

$$u_N(x) = N \sqrt{2 \ln 2} - \frac{1}{2} \ln(N \ln 2) + \ln(4\pi) \sqrt{2 \ln 2} + x \sqrt{2 \ln 2}. \quad (4.2.4)$$

The fact that $\mathbb{E} M_N(x) = O(1)$ precludes that there are too many exceedances of the level $u_N(x)$, by Chebyshev’s inequality. But it may still be true that the probability that $M_N(x) > u_N(x)$ tends to zero. To show that this is not the case, one may want to control, e.g., the variance of $M_N(x)$. Namely, using the Cauchy-Schwartz inequality, we have that

$$(\mathbb{E} M_N(x))^2 = \mathbb{E} (M_N(x) \mathbb{1}_{M_N(x) \geq 1})^2 \leq \mathbb{E} (M_N(x))^2 \mathbb{P}(M_N(x) \geq 1), \quad (4.2.5)$$

so that

$$\mathbb{P}(M_N(x) \geq 1) \geq \frac{(\mathbb{E} M_N(x))^2}{\mathbb{E} (M_N(x))^2}. \quad (4.2.6)$$

Now it is easy to compute

$$\mathbb{E} (M_N(x))^2 = 2^N \mathbb{P}(X_\sigma > u_N(x)) + 2^N (2^N - 1) \mathbb{P}(X_\sigma > u_N(x))^2 \quad (4.2.7)$$

$$= (\mathbb{E} M_N(x))^2 (1 - 2^{-N}) + \mathbb{E} M_N(x).$$

Hence

$$\mathbb{P}(M_N(x) \geq 1) \geq \frac{1}{1 - 2^{-N} + 1/\mathbb{E} M_N(x)} \sim \frac{1}{1 + e^x}, \quad (4.2.8)$$

for $N$ large. Together with the Chebyshev upper bound, this gives already good control on the probability to exceed $u_N(x)$, at least for large $x$:

$$\frac{e^{-x}}{1 + e^{-x}} \leq \mathbb{P} \left( \max_{\sigma \in S_N} X_\sigma > u_N(x) \right) \leq e^{-x}. \quad (4.2.9)$$

In particular,

$$\lim_{x \to \infty} \lim_{N \to \infty} \mathbb{P} \left( \max_{\sigma \in S_N} X_\sigma > u_N(x) \right) = 1, \quad (4.2.10)$$

which is a bound on the upper tail of the distribution of the maximum. This is called the second moment estimate. First and second moment estimates are usually easy to use and give the desired control on maxima, if they work. Unfortunately, they do not always work as nicely as here.
4.2.2 Law of the maximum and the extremal process

In this simple case, we can of course do better. In fact, the result on extremes of iid random variables from Chapters 1 and 2 imply that

$$P \left( \max_{\sigma \in \mathcal{S}_N} X_\sigma \leq u_N(x) \right) \to e^{-e^{-x}}, \quad (4.2.11)$$

as $N \uparrow \infty$. Moreover, with

$$\mathcal{E}_N \equiv \sum_{\sigma \in \mathcal{S}_N} \delta_{u_N^{-1}(X_\sigma)} \quad (4.2.12)$$

**Corollary 4.1.** In the REM, the sequence of point processes $\mathcal{E}_N$ converge in law to the Poisson point process with intensity measure $e^{-z}dz$ on $\mathbb{R}$.

4.3 The GREM, two levels

To understand what happens if correlation is introduced into the game, we consider the simplest version of the generalised random energy model with just two hierarchies. That is, for $\sigma_1 \in \mathcal{S}_{N/2}$ and $\sigma_2 \in \mathcal{S}_{N/2}$, we define the Hamiltonian

$$H_N(\sigma_1, \sigma_2) \equiv a_1 X_{\sigma_1}^1 + a_2 X_{\sigma_2}^2. \quad (4.3.1)$$

where all $X$ are iid centred Gaussian with variance $N$, and $a_1^2 + a_2^2 = 1$.

Note that this corresponds to the Gaussian process with covariance

$$\mathbb{E}H_N(\sigma)H_N(\sigma') = NA(q_N(\sigma, \sigma')), \quad (4.3.2)$$

where

$$A(x) = \begin{cases} 
0, & \text{if } x < 1/2 \\
a_1^2, & \text{if } 1/2 \leq x < 1, \\
1, & \text{if } x = 1. 
\end{cases} \quad (4.3.3)$$

4.3.0.1 Second moment method

We may be tempted to retry the second moment method that worked so nicely in the REM. Of course we get

$$\mathbb{E}M_N(x) = 2^N P(H_N(\sigma) > u_N(x)), \quad (4.3.4)$$

as in the REM. But this quantity does not see any correlation, so this should make us suspicious. But we know how to check whether this is significant: compute the second moment. This is easily done:
\[ E[M_N(x)^2] = \sum_{\sigma^1, \sigma^2, \tau^1, \tau^2} E \left[ I_{H_N(\sigma^1 \sigma^2) > u_N(x)} I_{H_N(\tau^1 \tau^2) > u_N(x)} \right] \]  \hspace{1cm} (4.3.5)

\[ = \sum_{\sigma^1, \sigma^2} E \left[ I_{H_N(\sigma^1 \sigma^2) > u_N(x)} \right] \]

\[ + \sum_{\sigma^1} \sum_{\sigma^2 \neq \tau^2} E \left[ I_{H_N(\sigma^1 \sigma^2) > u_N(x)} I_{H_N(\sigma^1 \tau^2) > u_N(x)} \right] \]

\[ + \sum_{\sigma^1 \neq \tau^1} \sum_{\sigma^2 \neq \tau^2} E \left[ I_{H_N(\sigma^1 \sigma^2) > u_N(x)} \right] E \left[ I_{H_N(\tau^1 \tau^2) > u_N(x)} \right]. \]

The first terms yield \(2^N \mathbb{P}(H_N(\sigma) > u_N(x))\), the last give \((2^N - 1)^2 \mathbb{P}(H_N(\sigma) > u_N(x))^2\), so we already know that these are ok, i.e. of order one. But the middle terms are different. In fact, a straightforward Gaussian computation shows that

\[ E \left[ I_{H_N(\sigma^1 \sigma^2) > u_N(x)} I_{H_N(\tau^1 \tau^2) > u_N(x)} \right] \sim \exp \left( -\frac{u_N(x)^2}{N(1 + a_1^2)} \right). \]  \hspace{1cm} (4.3.6)

Thus, with \(u_N(x)\) as in the REM, the middle term gives a contribution of order

\[ 2^{3N/2} - 2^{N/(1 + a_1^2)}. \]  \hspace{1cm} (4.3.7)

This does not explode only if \(a_1^2 \leq \frac{1}{2}\). What is going on here??

To see this, look in detail into the computations:

\[ \mathbb{P}(H_N(\sigma^1 \sigma^2) \sim u \wedge H_N(\sigma^1 \tau^2) \sim u) \sim \int \exp \left( -\frac{x^2}{2a_1^2 N} - \frac{(u - x)^2}{(1 - a_1^2)N} \right) dx, \]  \hspace{1cm} (4.3.8)

which is obtained by first integrating over the value \(x\) of \(X_{\sigma_1}^{\sigma_2}\), and then asking that both \(X_{\sigma_1 \tau_1}^{\sigma_2}\) and \(X_{\sigma_1 \tau_2}^{\tau_2}\) fill the gap to \(u\). Now this integral, for large \(u\) gets its main contribution from values of \(x\) that maximise the exponent, which is easily checked to be reached at \(x_c = \frac{2a_1^2}{1 + a_1^2} \). For \(u = u_N(x)\), this yields indeed (??). However, in this case

\[ x_c \sim N \frac{2a_1^2 \sqrt{2 \ln 2}}{1 + a_1^2} = \frac{2 \sqrt{2} a_1}{1 + a_1^2} a_1 N \sqrt{\ln 2}. \]  \hspace{1cm} (4.3.9)

Now \(a_1 N \sqrt{\ln 2}\) is the maximum that any of the \(2^{N/2}\) variables \(a_1 X_{\sigma_1}^{\sigma_2}\) can reach, and the value \(x_c\) is larger than that as soon as \(a_1 > \sqrt{2} - 1\). This indicates that this moment computation is non-sensical above that value of \(a_1\). On the other hand, when we compute \(\mathbb{P}(H_N(\sigma^1 \sigma^2) \sim u_N)\) in the same way, we find that the corresponding critical value of the first variable is \(x_c = u_N a_1^2 = a_1 \sqrt{2} a_1 N \sqrt{\ln 2} = a_1 \sqrt{2} \max_{\sigma_1} (a_1 X_{\sigma_1}^{\sigma_2}).\) Thus, here a problem occurs only for \(a_1 > 1/\sqrt{2}\). In that latter case we must expect a change in the value of the maximum, but for smaller values of \(a_1\) we should just be more clever.
4.3 The GREM, two levels

4.3.0.2 Truncated second moments

We see that the problem comes with the excessively large values of the first component contributing to the second moments. A natural idea is to cut these off. Thus we introduce

\[ \hat{M}_N(x) \equiv \sum_{\sigma^1, \sigma^2} \mathbb{I}_{H_N(\sigma^1 \sigma^2) > u_N(x)} \mathbb{I}_{a_1 X_{\sigma^1}^1 \leq a_1 N \sqrt{\ln 2}}. \]  \hspace{1cm} (4.3.10)

A simple computation shows that

\[ \mathbb{E} \hat{M}_N(x) = \mathbb{E} M_N(x) (1 + o(1)), \]  \hspace{1cm} (4.3.11)

as long as \( a_1^2 < 1/2 \). On the other hand, when we now compute \( \mathbb{E} [\hat{M}_N(x)^2] \), the previously annoying term becomes

\[ 2^{3N/2} \mathbb{E} \left[ \mathbb{I}_{H_N(\sigma^1 \sigma^2) > u_N(x)} \mathbb{I}_{H_N(\sigma^1 \tau^2) > u_N(x)} \mathbb{I}_{X_{\sigma^1}^1 \leq a_1 N \sqrt{\ln 2}} \right] \sim 2^{-N (1 - \sqrt{2/a_1})^2 / (1 - a_1^2)}, \]  \hspace{1cm} (4.3.12)

which is insignificant for \( a_1^2 < 1/2 \). Since \( M_N(x) \geq \hat{M}_N(x) \), it follows that

\[ \mathbb{E} M_N(x) \geq \mathbb{P} (M_N(x) \geq 1) \geq \mathbb{P} (\hat{M}_N(x) \geq 1) \geq \left( \frac{\mathbb{E} \hat{M}_N(x)^2}{\mathbb{E} M_N(x)} \right)^2 = O(1). \]  \hspace{1cm} (4.3.13)

So the result can be used to show that, as long as \( a_1^2 < 1/2 \), the maximum is of the same order as in the REM. If \( a_1^2 = 1/2 \), the bound on (4.3.12) is order one. Thus we still get the same behaviour for the maximum. If \( a_1 \) is even larger, then the behaviour of the maximum changes. Namely, one cannot reach the value \( \sqrt{2 \ln 2} N \) any more, and the maximum will be achieved by adding the maxima of the first hierarchy to the maximum in one of the branches of the second hierarchy. This yield for the leading order \( (a_1 + a_2) \sqrt{2N} \).

4.3.1 The finer computations

As in the REM, we do of course want to compute things more precisely. This means we want to compute the limit of

- \( \mathbb{P} \left( \max_{(\sigma^1, \sigma^2) \in S_{2N/2}} H_N(\sigma^1 \sigma^2) \leq u_N(x) \right) \), or, better,
- compute the limit of the Laplace functionals of the extremal process,

\[ \mathbb{E} \exp \left( - \sum_{\sigma^1, \sigma^2 \in S_{N/2}} \phi \left( u_N^{-1} (H_N(\sigma^1 \sigma^2)) \right) \right). \]  \hspace{1cm} (4.3.14)
Technically, there is rather little difference in the two cases, let us for simplicity compute the Laplace functional for the case $\phi(x) = \lambda \mathbb{I}_{u_N(x) = 0}$. Then

$$
E \left[ \exp \left( -\lambda \sum_{\sigma^1, \sigma^2 \in \mathcal{S}_N} \mathbb{I}_{H_N(\sigma^1, \sigma^2) > u_N(x)} \right) \right] \tag{4.3.15}
$$

$$
= \prod_{\sigma^1 \in \mathcal{S}_N} E \left[ \exp \left( -\lambda \sum_{\sigma^2 \in \mathcal{S}_N} \mathbb{I}_{H_N(\sigma^1, \sigma^2) > u_N(x)} \right) \right]
$$

$$
= \left[ \frac{1}{\sqrt{2\pi N}} \int_{-\infty}^{\infty} e^{-t^2/2N} \left( 1 + (e^{-\lambda} - 1) P(a_2 X > u_N(x) - a_1 t) \right)^2 \right]^{N/2}
$$

Here $X$ is a centred Gaussian random variable with variance $N$.

Basically, to get something nontrivial, the probability in the last expression should be at most order $2^{-N}$. Namely, if this probability is much larger than $2^{-N}$, then

$$
\left( 1 + (e^{-\lambda} - 1) P(a_2 X > u_N(x) - a_1 t) \right)^2 \leq e^{-2^{-N/2}}. \tag{4.3.16}
$$

Integrating this over $t$ does not make it any bigger. So the last line in (4.3.15) is of the form

$$
\text{Biggl[} \frac{1}{\sqrt{2\pi N}} \int_{-\infty}^{t_c} e^{-t^2/2N} \left( 1 + (e^{-\lambda} - 1) P(a_2 X > u_N(x) - a_1 t) \right)^2 dt \right]^{N/2}
$$

$$
+ \frac{1}{\sqrt{2\pi N}} \int_{t_c}^{\infty} e^{-t^2/2N} O \left( e^{-2^{-N/2}} \right) \right]^{N/2}. \tag{4.3.17}
$$

Here we defined $t_c$ by

$$
P(a_2 X > u_N(x) - a_1 t) = 2^{-N}. \tag{4.3.18}
$$

By standard Gaussian asymptotics, this yields

$$
t_c = \frac{u_N(x)}{a_1} - \frac{a_2 N \sqrt{\ln 2}}{a_1}. \tag{4.3.19}
$$

The first term in (4.3.17) is essentially

$$
\frac{1}{\sqrt{2\pi N}} \int_{-\infty}^{t_c} e^{-t^2/2N} \left( 1 + 2^{N/2} (e^{-\lambda} - 1) P(a_2 X > u_N(x) - a_1 t) \right) dt. \tag{4.3.20}
$$

Inserting the standard Gaussian asymptotics and reorganising things, this is asymptotically equal to
4.4 The general model and relation to branching Brownian motion

We have seen that the general case corresponds to a Gaussian process indexed by $\mathcal{S}_N$ with covariance $\mathbb{E}X_\sigma X_{\sigma'} = NA(q_N(\sigma, \sigma'))$, for a non-decreasing function $A$. Now we are invited to think of $\mathcal{S}_N$ as the leaves of a tree with binary branching and $N$ levels. Then it makes sense to extend the Gaussian process from the leaves of the tree to the entire tree. If we think of the edges of the tree of being of length one, this process should be indexed by $t \in [0, N]$, with covariance

$$\mathbf{1} + \left( e^{-\lambda} - 1 \right) \frac{2N/2}{\sqrt{2\pi u_N(x)/\sqrt{N}}} \int_{-\infty}^{t_c - u_N(x) \alpha} e^{-\frac{s^2}{2N}} ds$$

$$\sim 1 + 2^{-N/2} e^{-x} \left( e^{-\lambda} - 1 \right) \int_{-\infty}^{t_c - u_N(x) \alpha} e^{-\frac{s^2}{2N}} ds. \quad (4.3.21)$$

The integral is essentially one, if $t_c - u_N(x) \alpha > 0$, which is the case when $\alpha^2 < 1/2$. In the case $\alpha^2 = 1/2$, the integral is 1/2. The second term in (4.3.17) is bounded by

$$O \left( e^{-2^{-N/2}} \right) e^{-t_c^2/2N}. \quad (4.3.22)$$

Hence this term gives a vanishing contribution to (4.3.17), and we find that we get in both cases

$$\lim_{N \to \infty} \mathbb{E} \exp \left( -\frac{\lambda}{\sigma^1, \sigma^2 \in \mathcal{S}_N} \mathbf{1}_{H_N(\sigma^1 \sigma^2) > u_N(x)} \right)$$

$$= \exp \left( e^{-x} \left( e^{-\lambda} - 1 \right) K \right), \quad (4.3.23)$$

which is the Laplace functional for a Poisson process with intensity $Ke^{-x} dx$, $K$ being 1 or 1/2, depending on whether $\alpha^2 < 1/2$ or $\alpha^2 = 1/2$, respectively.

In [19] the full picture is explored when the number of levels is arbitrary (but finite). The general result is that the extremal process remains Poisson with intensity $Ke^{-x} dx$, whenever the function $A$ is strictly below the straight line $A(x) = x$, and it is Poisson with intensity $Ke^{-x} dx$ when $A(x)$ touches the straight line finitely many times. The value of the constant $K < 1$ can be expressed in terms of the probability that a Brownian bridge stays below 0 in the points where $A(x) = x$.

If $\alpha^2 > 1/2$, the entire picture changes. In that case, the maximal values of the process are achieved by adding up the maxima of the two hierarchies. and this leads to a lower value even on the leading scale $N$. The extremal process also changes, but it is simply a concatenation of Poisson processes. This has all been fully explored in [19].
in construction

\[ E[X_\sigma(t)X_{\sigma'}(s)] \equiv NA((t \land s)/N \land q_N(\sigma, \sigma')). \quad (4.4.1) \]

These models were introduced by Derrida and Gardner \[38, 39\] and further investigated in \[20\]. It turns out that the leading order behaviour of the maximum in these models can be computed by approximating the function \(A\) by step functions. The point here is that if \(A\) is a step functions, one can show easily, using the methods explained above, that

\[ \lim_{N \to \infty} N^{-1} \max_{\sigma \in S_N} X_\sigma = \sqrt{2 \ln 2} \int_0^1 \sqrt{\bar{A}(x)} dx, \quad (4.4.2) \]

where \(\bar{A}\) denotes the convex envelope of \(A\) (i.e. the smallest convex function larger than or equal to \(A\)). It is straightforward to see that this formula remains true for arbitrary non-decreasing \(A\). The proof, given in \[20\], uses Gaussian comparison method explained in Chapter 3.

A similar feature is not true for the subleading corrections. They can be computed for all step functions, but the expression has no obvious limit then these converge to some continuous function. If the convex hull of \(A\) is the linear function \(x\) we have seen already in the case of two steps that the intensity of the Poisson process of extremes changes when \(A\) touches the straight line. In general, a rather complicated looking formula is obtained if \(A\) touches the straight line in several points \[20\]. These observations make the case \(A(x) = x\) an obviously interesting target.

Another interesting observation is that the processes introduced above can all be constructed with the help of Brownian motions. In the case \(A(x) = x\), one sees easily that this can be realised as follows. Start at time 0 with a Brownian motion that splits into two independent Brownian motions at time \(t = 1\). Each of these splits again into two independent copies at time \(t = 2\), and so on. The case \(A\) non-decreasing just corresponds to a time change of this process. The case studied above, when \(A\) is a step function, corresponds to singular time changes where for some time the Brownian motions stop to move and just split and then recover their appropriate size instantaneously.

### 4.5 The Galton-Watson process

So far we have looked at Gaussian processes on \(\{-1, 1\}^N\) which could be seen as the leaves of a binary tree of depth \(N\). We could also decide to take some different tree. A relevant example would be the continuous time Galton-Watson tree. See, e.e. \[8\].

The Galton-Watson process \[74\] is the most canonical continuous time branching process. Start with a single particle at time zero. After an exponential time of parameter one, this particle splits into \(k\) particles according to some probability distribution \(p\) on \(\mathbb{N}\). Then each of the new particles splits at independent exponential
4.5 The Galton-Watson process

In construction times independently according to the same branching rule, and so on. For the purposes of this book, we will always assume that \( p_0 = 0 \), i.e. no deaths occur.

At time \( t \), there are \( n(t) \) "individuals" \( (i_k(t), 1 \leq k \leq n(t)) \). The point is that the collection of individuals is endowed with a genealogical structure. Namely for each pair of individuals at time \( t \), \( i_k(t), i_r(t) \), there is a unique time \( 0 \leq s \leq t \) when they shared a common ancestor for the last time. We call this time \( d(i_k(t), i_r(t)) \).

It is sometimes useful to provide a labelling of the Galton-Watson process in terms of multi-indices. For convenience we think of multi-indices as infinite sequences of non-negative integers. Let us set

\[
I \equiv \mathbb{Z}^N_+ ,
\]

and let \( F \subset I \) denote the subset of multi-indices that contain only a finitely many entries that are different from zero. Ignoring leading zeros, we see that

\[
F = \bigcup_{k=0}^\infty \mathbb{Z}^k_+ ,
\]

where \( \mathbb{Z}^0_+ \) is either the empty multi-index or the multi-index containing only zeros.

A continuous-time Galton-Watson process will be encoded by the set of branching times, \( \{ t_1 < t_2 < \cdots < t_{\ell(0)} < \cdots \} \) (where \( \ell(0) \) denotes the number of branching times up to time \( t \)) and by a consistently assigned set of multi-indices for all times \( t \geq 0 \). To do so, we construct for a given tree the sets of multi-indices, \( \tau(t) \) at time \( t \) as follows.

- \( \tau(0) = \{ (0,0, \ldots) \} \).
- for all \( j \geq 0 \), for \( t \in [t_j, t_{j+1}) \), for all \( \tau(t) = \tau(t_j) \).
- If \( i \in \tau(t_j) \) then \( i + (0, \ldots, 0, k, 0, \ldots) \in \tau(t_{j+1}) \) if \( 0 \leq k \leq \ell_i(t_{j+1}) - 1 \), where\[ \ell_i(t_j) = \# \{ \text{offsprings of the particle corresponding to } i \text{ at time } t_j \}. \]

Note that here we use the convention that, if a given branch of the tree does not "branch" at time \( t_j \), it is considered as having one offspring.

We can relate the assignment of labels in a backwards consistent fashion as follows. It \( i \equiv (i_1, i_2, i_3, \ldots) \in \tau(t) \), we define the function \( i(t) \), for \( r \leq t \), of multi-indices via

\[
i_i(r) = \begin{cases} i_t, & \text{if } t_r \leq r, \\ 0, & \text{if } t_r > r. \end{cases}
\]

Clearly, \( i(r) \in \tau(r) \). This construction allows to define the boundary of the tree at infinite time as follows:

\[
\partial T \equiv \{ i \in I : \forall t < \infty, i(t) \in \tau(t) \} .
\]

Note that \( \partial T \) is an ultrametric space equipped with the ultrametric \( m(i, j) \equiv e^{-d(i, j)} \), where \( d(i, j) = \sup \{ t \geq 0 : i(t) = j(t) \} \).
Note that from the knowledge of the set of multi indices in $\partial T$ and the set of branching times, the entire tree can be reconstructed. Similarly, knowing $\tau(t)$ allows to reconstruct the tree up to time $t$.

**Remark 4.2.** The labelling of the GW-tree is a slight variant of the familiar Ulam-Neveu-Harris labelling (see e.g. [41]). In our labelling the added zeros keep track of the order in which branching occurred in continuous time. The construction above is also nicely explained for discrete time GW processes in [73].

![Fig. 4.1 Construction of $\tilde{T}$: The green nodes were introduced into the tree 'by hand'.](image)

### 4.6 The REM on the Galton-Watson tree

A Gaussian process can then be constructed where the covariance is a function of this distance, just like on the binary tree. The simplest Gaussian process on such trees is of course the REM. i.e. at time $t$ where are $n(t)$ iid independent Gaussian random variables $x_k(t), k = 1, \ldots, n(t)$, of variance $t$. We will always be interested in the case then $\sum_k kp_k > 1$, i.e. when the branching process is supercritical. In that case $\lim_{t \to \infty} t^{-1} \ln n(t) = c > 1$, and $\lim n(t)/E n(t) = e$, where $e$ is an exponential random variable with parameter 1. We will discuss more in that in the next chapter. What does that mean for the REM? If we condition on $n(t)$, Our standard estimates will apply and we get that
\[ P \left( \max_{k \leq n(t)} x_k(t) \leq u_n(x(t)) \right) \sim e^{-e^{-x}}, \quad (4.6.1) \]

provided
\[ P (x_k(t) > u_n(x(t))) \sim 1/n. \quad (4.6.2) \]

Using the computations in Chapter 1 and taking into account that the variance of \( x_k(t) \) is equal to \( t \), this implies that
\[ u_n(x(t)) = t \sqrt{2t^{-1} \ln n - \frac{\ln \ln n + \ln(4\pi)}{2} + \frac{x}{\sqrt{2t^{-1} \ln n}}}. \quad (4.6.3) \]

Passing to the limit, since the right hand side of (4.6.1) does not depend on \( n \), one easily sees that
\[ P \left( \max_{k \leq n(t)} x_k(t) \leq u_{n(t)}(x(t)) \right) \rightarrow e^{-e^{-x}}. \quad (4.6.4) \]

Taking into account the convergence properties of \( n(t) \), we see that, asymptotically as \( t \uparrow \infty \),
\[ u_{n(t)}(x(t)) \sim t \sqrt{2c} - \frac{\ln c + \ln(4\pi) + \ln(x)}{\sqrt{2c}} + \frac{x}{\sqrt{2c}}. \quad (4.6.5) \]

From this we can see that instead of the random rescaling, we can also use a deterministic rescaling
\[ u_t(x) \equiv t \sqrt{2c} - \frac{\ln c + \ln(4\pi) + \ln(x)}{\sqrt{2c}} + \frac{x}{\sqrt{2c}}. \quad (4.6.6) \]

In that case we get that
\[ P \left( \max_{k \leq n(t)} x_k(t) \leq u_t(x) \right) \rightarrow \mathbb{E} \left[ \exp \left( -\frac{e}{4\pi} e^{-\sqrt{2c} x} \right) \right], \quad (4.6.7) \]

where the average is over the exponential random variable \( e \). Here we see for the first time the appearance of a random shift of a Gumbel distribution. Note that I moved some other constants around to arrive at a form for the right hand side that is conventional in the context of branching Brownian motion.

In the same way we see that for the REM on the Galton-Watson tree, the extremal process acquires a random element if we chose the same deterministic rescaling. The processes
\[ \sum_{k=1}^{n(t)} \delta_{u_t^{-1}(x_k(t))} \rightarrow \mathbb{E} \left[ \text{PPP} \left( \frac{e}{4\pi} e^{-\sqrt{2c} x} dx \right) \right]. \quad (4.6.8) \]

In other words, the limiting process is a Poisson process with random intensity, also known as a Cox process.
Chapter 5
Branching Brownian motion

In this chapter we start to look at branching Brownian motion as a continuous time version of the GREM. We collect some basic facts that will be needed later.

5.1 Definition and basics

The simplest way to describe branching Brownian motion is as follows. At time zero, a single particle \( x_1(0) \) starting at the origin, say, begins to perform Brownian motion in \( \mathbb{R} \). After an exponential time, \( \tau \), of parameter one, the particle splits into two identical, independent copies of itself that start Brownian motion at \( x_1(\tau) \). This process is repeated ad infinitum, producing a collection of \( n(t) \) particles \( x_k(t), 1 \leq k \leq n(t) \).

This construction can be easily extended to the case of more general offspring distributions where particles split into \( k \geq 1 \) particles with probabilities \( p_k \). We will always assume that \( \sum_{k \geq 1} p_k = 1 \), \( \sum_{k \geq 1} kp_k = 2 \) and \( \sum_{k \geq 2} k(k-1)p_k < \infty \). Note that while of course nothing happens if the particle splits into one particle, we keep this option to be able to maintain the mean number of offspring equal to 2.

Another way of constructing BBM is to first build a Galton-Watson tree with the same branching mechanism as above. Branching Brownian motion (at time \( t \)) can then be interpreted as a Gaussian process \( x_k(t) \), indexed by the leaves of a GW-process, such that \( \mathbb{E}x_k(t) = 0 \) and, given a realisation of the Galton Watson process,

\[
\mathbb{E}[x_k(t)x_\ell(t) = d(i_k(t), i_\ell(t))].
\] (5.1.1)

This observation makes BBM perfectly analogous to the CREM [20] with covariance function \( A(x) = x \), which we have seen to be the critical covariance.

57
5.2 Rough heuristics

Already in the GREM we have seen that than $A$ touches the straight line $x$, there appears a constant in the intensity measure that reflects the fact that the process is not allowed to get close to the maximal values possible at this point. The basic intuition of what happens in BBM is that the ancestral paths of BBM have to stay below the straight line with slope $\sqrt{2}$ all the time. This will force the maximum down slightly. The rough heuristics of the correct rescaling can be obtained from second moment calculations. More precisely, we should consider the events

$$\left\{ \exists k \leq n(t) : x_k(t) - m(t) > x \right\} \wedge \left\{ x_k(s) \leq \sqrt{2}s, \forall s \in (r, t-r) \right\}, \quad (5.2.1)$$

for $r$ fixed. Here we write, for given $t$, and $s \leq t$, $x_k(s)$ for the ancestor at time $s$ of the particle $x_k(t)$. One of the observations of Bramson [26] is that the event in (5.2.1) has the same probability as the event

$$\left\{ \exists k \leq n(t) : x_k(t) - m(t) > x \right\}, \quad (5.2.2)$$

when first $t \uparrow \infty$ and then $r \uparrow \infty$. This is not very easy, but also not too hard. Let us first recall that a Brownian bridge from $0$ to $0$ in time $t$ can be represented as

$$z(s) = B(s) - \frac{s}{t}B(t). \quad (5.2.3)$$

More generally, we will denote by $z_{x,y}^t$ a Brownian bridge from $x$ to $y$ in time zero. Clearly,

$$z_{x,y}^t(s) = x + (y-x)\frac{s}{t} + z_{0,0}^t(s). \quad (5.2.4)$$

It is clear that the ancestral path from $x_k(t)$ to zero is just a Brownian bridge from $x_k(t)$ to zero. That is, we can write

$$x_k(t-s) = \frac{t-s}{t} x_k(t) + z_{k,t}^s,$$  

where $z_{k,t}^s, s \in [0, t]$, is a Brownian bridge form $0$ to $0$, independent of $x_k(t)$. The event in (5.2.1) can then be expressed as

$$\left\{ \exists k \leq n(t) : x_k(t) - m(t) > x \right\} \wedge \left\{ z_{k,t}^s(s) \leq (1-s/t)(\sqrt{2}s - x_k(t)), \forall s \in (r, t-r) \right\}, \quad (5.2.6)$$

where of course the Brownian bridges are not independent. The point, however, is, that as far as the maximum is concerned, they might as well be independent.

We need to bound the probability of a Brownian bridge to stay below a straight line. The following lemma is taken from Bramson [22], but is of course classic.

**Lemma 5.1.** Let $y_1, y_2 > 0$. Let $z$ be a Brownian bridge from $0$ to $0$ in time $t$. Then

$$\mathbb{P} \left( z(s) \leq \frac{1}{t}y_1 + \frac{1}{r}y_2, \forall 0 \leq s \leq t \right) = 1 - \exp(-y_1y_2/t). \quad (5.2.7)$$
5.2 Rough heuristics

The proof and further discussions on properties of Brownian bridges will be given in the next chapter.

The first moment estimate shows (the calculations are a bit intricate and constitute a good part of Bramson’s first paper [26])

\[
\mathbb{P} \left( \exists k \leq n(t) : \left\{ x_k(t) - m(t) > x \right\} \wedge \left\{ \beta_k(s) \leq (1 - s/t)(\sqrt{2}s - x_k(t)) \right\}, \forall s \in (r, t - r) \right) \\
\sim e^{- \frac{(m(t)+x)^2}{2}} \frac{C \chi^2}{t} \sim C'e^{- \frac{(m(t)+x)^2}{2t^2}} x^2 e^{- \sqrt{2}x}.
\]  

(5.2.8)

The t-dependent term is equal to one if

\[
m(t) = \frac{3}{2\sqrt{2}} \ln t.
\]  

(5.2.9)

To justify that this upper bound is not bad, one has to perform a second moment computation in which one retains the condition that the particle is not above the straight line \(\sqrt{2}s\). This is not so easy here and actually yields a lower bound there the \(x^2\) is not present. Indeed, it also turns out that the upper bound is not optimal, and that the \(x^2\) should be replaced by \(x\). We will show that in the next chapter.

A refinement of the truncated second moment method was used more recently by Roberts [67].

One can actually localise the ancestral paths of extremal particles much more precisely, as was shown in [3]. Namely, define

\[
f_t, \alpha(s) = \begin{cases} 
s^\alpha, & 0 \leq s \leq t/2 \\
(t - s)^\alpha, & t/2 \leq s \leq t.
\end{cases}
\]  

(5.2.10)

Then the following holds:

**Theorem 5.2 ([3]).** Let \(D = [d_l, d_u] \in \mathbb{R}\) be a bounded interval. Then for any \(0 < \alpha < \frac{1}{2} < \beta\), and any \(\varepsilon > 0\), there are \(r_0 \equiv r_0(D, \varepsilon, \alpha, \beta)\), such that for all \(r > r_0\) and \(t > 3r\),

\[
\mathbb{P} \left( \forall k \leq n(t) : s.t. x_k(t) - m(t) \in D, \text{ it holds that} \right)
\]

\[
\forall s \in [t - r, t], \frac{s}{t} x_k(t) + f_{t, \alpha}(s) \leq x_k(s) \leq \frac{s}{t} x_k(t) + f_{t, \alpha}(s) \geq 1 - \varepsilon.
\]  

(5.2.11)

Basically this theorem says that the paths of maximal particles lie below the straight line to their endpoint along the arc \(f_{t, \alpha}(s)\). This is just due to a property of Brownian bridges: in principle, a Brownian bridge wants to oscillate (a little bit more) than \(\sqrt{t}\) at its middle. But if it is not allowed to cross zero, then it also will not get close to zero, because each time it does so, it is unlikely to not do the crossing. This phenomenon is known as *entropic repulsion* in statistical mechanics.

Note also that for the ancestral paths of particles that end up much lower than \(\sqrt{2}t\), say around \(at\) with \(a < \sqrt{2}\), there is no such effect present. The ancestral paths
of these particles are just Brownian bridges from zero to their end position which are localised around the straight line by a positive upper and a negative lower envelope $f_{\gamma t}$, with any $\gamma > \frac{1}{2}$.

5.3 Recursion relations

The presence of the tree structure in BBM of course suggest to derive recursion relations. To get into the mood, let us show that

**Lemma 5.3.**

\[ \mathbb{E} n(t) = e^t. \]  

**Proof.** For notational simplicity we consider only the case of pure binary branching. The idea is to use the recursive structure of the process. Let $\tau$ be the first branching time. Clearly, if $t < \tau$, then $n(t) = 1$. Otherwise, it is $n'(t-\tau) + n''(t-\tau)$, where $n'$ and $n''$ are the number of particles in the two independent offspring processes. This reasoning leads to

\[ \mathbb{E} n(t) = \mathbb{P}(\tau > t) + \int_0^t \mathbb{P}(\tau \in ds) 2 \mathbb{E} n(t-s) \]  

\[ = e^{-t} + 2 \int_0^t e^{-s} \mathbb{E} n(t-s) ds. \]

Differentiating this equation yields

\[ \frac{d}{dt} \mathbb{E} n(t) = -e^{-t} + 2e^{-t} \mathbb{E} n(0) + 2 \int_0^t e^{-s} \frac{d}{ds} \mathbb{E} n(t-s) ds \]  

\[ = -e^{-t} + 2e^{-t} \mathbb{E} n(0) - 2 \int_0^t e^{-s} \frac{d}{ds} \mathbb{E} n(t-s) ds \]  

\[ = -e^{-t} + 2 \mathbb{E} n(t) - 2 \int_0^t e^{-s} \mathbb{E} n(t-s) ds = \mathbb{E} n(t), \]

where we used integration by parts in the second inequality and the fact that $n(0) = 1$. The assertion follows by solving this differential equation with $\mathbb{E} n(0) = 1$. \qed

We can also show the following classical result (see the standard textbook by Athreya and Ney [8] for this and many further results on branching processes):

**Lemma 5.4.** If $n(t)$ is the number of particles of BBM at time $t$, then

\[ M(t) \equiv e^{-t} n(t) \]  

is a martingale. Moreover, $M(t)$ converges, a.s. and in $L^1$, to an exponential random variable of parameter 1.
5.4 The F-KPP equation

Proof. Again we write things only for the binary case. The verification that $M(t)$ is a martingale is elementary. Since $M(t)$ is positive with mean one, it is bounded in $L^1$ and hence, by Doob’s martingale convergence theorem, $M(t)$ converges a.s. to a random variable $M$. To show that the martingale is uniformly integrable and hence converges in $L^1$, we show that

$$\phi(t) \equiv \mathbb{E}[M(t)^2] = 2 - e^{-t}. \quad (5.3.5)$$

This can be done by noting that $\mathbb{E}M(t)^2$ satisfies the recursion

$$\phi(t) = e^{-3t} + 2 \int_0^t e^{-3s} \phi(t-s)ds + \frac{2}{3} (1 - e^{-3t}). \quad (5.3.6)$$

Differentiating yields the differential equation

$$\phi'(t) = 2 - \phi(t), \quad (5.3.7)$$

of which (5.3.5) is the unique solution with $\phi(0) = 1$. Convergence to the exponential distribution can be proven similarly. ⊓ ⊔

5.4 The F-KPP equation

The fundamental link between BBM and the F-KPP equation [54, 37] is generally attributed to McKean [61], but appears already in Skorohod [68] and Ikeda, Naga-sawa, and Watanabe [44, 45, 46].

Lemma 5.5 ([61]). Let $f: \mathbb{R} \to [0, 1]$ and $\{x_k(t): k \leq n(t)\}$ a branching Brownian motion starting at 0. Let

$$v(t,x) \equiv \mathbb{E}\left[\prod_{k=1}^{n(t)} f(x-x_k(t))\right]. \quad (5.4.1)$$

Then, $u(t,x) \equiv 1 - v(t,x)$ is the solution of the F-KPP equation (5.4.2)

$$\partial_t u = \frac{1}{2} \partial_x^2 u + F(u), \quad (5.4.2)$$

with initial condition $u(0,x) = 1 - f(x)$ where $F(u) = (1-u) - \sum_{k=1}^\infty p_k(1-u)u_k$.

Remark 5.6. The reader may wonder why we introduce the function $v$. It is easy to check that $v(t,x)$ itself solves equation (5.4.2) with $F(u)$ replaced by $-F(1-v)$. The form of $F$ given in the lemma is the one customary known as F-KPP equation and will be used throughout this text.

Proof. The derivation of the F-KPP equation is quite similar to the arguments used in the previous section. Again we restrict to the case of binary branching. Let $f:
\( \mathbb{R} \to [0, 1] \). Define \( v(t, x) \) by (5.4.1). Then, distinguishing the cases when the first branching occurs before or after \( t \), we get

\[
v(t, x) = e^{-t} \int e^{-\frac{z^2}{2t}} f(x - z) dz + \int_0^t e^{-s} \int e^{-\frac{z^2}{2s}} v(t - s, x - z)^2 dz ds.
\]

Differentiating with respect to \( t \), using integration by parts as before together with the fact that the heat kernel satisfies the heat equation, we find that

\[
\partial_t v = \frac{1}{2} \partial^2_x v + v^2 - v.
\]

Obviously, \( v(0, x) = f(x) \). Following the remark above, \( u = 1 - v \) then solves (5.4.2).

There is a smart way to see that (5.4.4) and (5.4.3) are the same thing. Namely, if \( H(t, x) \equiv e^{-t} \int e^{-\frac{z^2}{2t}} \), then is the Green kernel for the linear operator

\[
\partial_t - \frac{1}{2} \partial^2_x + 1,
\]

i.e. the solution of the inhomogeneous linear equation

\[
\partial_t v - \frac{1}{2} \partial^2_x v + v = r
\]

with initial condition \( u_0 \) is given by

\[
v(t, x) = \int H(t, x - y) v_0(y) dy + \int_0^t \int H(s, y) r(t - s, x - y) ds dy.
\]

Inserting \( r(t, x) = v^2(t, x) \), we obtain what is known as the mild formulation of Eq. (5.4.4), and this is precisely (5.4.3). \( \square \)

The first example is obtained by choosing \( f(x) = \mathbb{1}_{x \geq 0} \). Then

\[
v(t, x) = \mathbb{E} \left[ \prod_{k=1}^{n(t)} \mathbb{1}_{x - x_k(t) \geq 0} \right] = \mathbb{P} \left( \max_{k \leq n(t)} x_k(t) \leq x \right).
\]

A second example we will exploit is obtained by choosing \( f(x) = \exp(-\phi(x)) \) for \( \phi \) a non-negative continuous function with compact support. Then

\[
v(t, x) = \mathbb{E} \left[ \prod_{k=1}^{n(t)} e^{-\phi(x - x_k(t))} \right] = \mathbb{E} \left[ \exp(- \int \phi(x - z) \mathcal{E}_t(dz)) \right],
\]
5.5 The travelling wave

where $\delta_t \equiv \sum_{n=1}^{\sigma(t)} \delta_n(t)$ is the point process associated to BBM at time $t$. We see that in this case $u$ is the Laplace functional of the point process $\delta_t$.

**Definition 5.7.** We call the equation (5.4.2) with a general non-linear term the F-KPP equation. We say that $F$ satisfies the standard conditions if $F \in C^1([0,1])$, $F(0) = F(1) = 0$, $F(u) > 0, \forall u \in (0,1)$, $F'(0) = 1$, $F'(u) \leq 1, \forall u \in [0,1]$, and

\[ 1 - F'(u) = O(u^p). \quad (5.4.11) \]

5.5 The travelling wave

Bramson [22] studied convergence of solutions to a large class of KPP equations for a large class of functions $F$ that include the those that arise from BBM. In particular, he established under what conditions on the initial data solutions converge to travelling waves. Part of this is based on earlier results by Kolmogorov et al [54] and Uchiyama [72]. For a purely probabilistic analysis of the travelling wave solutions, see also Harris [43].

We will always be interested in initial conditions that lie between zero and one. It is easy to see that then the solutions remain between zero and one forever.

A travelling wave moving with speed $\lambda$ would be a solution, $u$, of the F-KPP equation such that

\[ \frac{d}{dt} u(t, x+\lambda t) = 0. \quad (5.5.1) \]

A simple computation shows that this implies that $u(t, x+\lambda t) = w_\lambda(x)$, where

\[ \frac{1}{2} \partial_x^2 w_\lambda(x) + \lambda \partial_x w_\lambda(x) + F(w_\lambda(x)) = 0. \quad (5.5.2) \]

If we want a solution that decays to zero at $+\infty$, for large positive $x$, $w$ must be close to the solution of the linear equation

\[ \frac{1}{2} \partial_x^2 w(x) + \lambda \partial_x w(x) + w(x) = 0 \quad (5.5.3) \]

But his equation can be solved explicitly. If $\lambda \neq \sqrt{2}$, then there are two linearly independent solutions $e^{-b_\pm x}$, where $b_\pm = \frac{\lambda}{2} \pm \sqrt{\lambda^2 - 2}$. Clearly, these values are real only if $\lambda > \sqrt{2}$, so only in this case we can have non-oscillatory solutions. If $\lambda = \sqrt{2}$, then there are two solutions $e^{-\sqrt{2}x}$ and $xe^{-\sqrt{2}x}$. Kolmogorov et al [54] and Uchiyama [72] showed by a phase-space analysis that in both cases the heavier tailed solution describes the correct asymptotics of the travelling wave solution, and that it is unique, up to translations.
Lemma 5.8 ([54][72]). Let $F$ satisfy the standard conditions. Then, for $\lambda \geq \sqrt{2}$, Equation (5.5.2) has a unique solution satisfying $0 < w_2(x) < 1$, $w_2(x) \to 0$, as $x \to +\infty$, and $w_2(x) \to 1$, as $x \to -\infty$, up to translation, i.e. if $w, w'$ are two solutions, then there exists $a \in \mathbb{R}$ s.t. $w'_\lambda(x) = w_\lambda(x + a)$.

Proof. Note first, that if $w_\lambda \in [0, 1]$, then, unless $w_\lambda$ is constant, it must hold that for all $x \in \mathbb{R}$, $w_\lambda(x) \in (0, 1)$. For, if for some $x \in \mathbb{R}$, $w_\lambda(x) = 0$, then it must also hold that $\partial_x w_\lambda(x) = 0$. But then, since $F(w_\lambda(x)) = 0$, the initial value problem with these initial data at $x$ has the unique solution $w_\lambda = 0$. The same holds if $w_\lambda(x) = 1$. Next we look at the equation in phase space. It reads

$$
\begin{align*}
q' &= p \\
p' &= -2F(q) - 2\lambda p.
\end{align*}
$$

Clearly, this has the two fix points $(0, 0)$ and $(1, 0)$. The Hessian matrices at these fixpoints are

$$
\begin{pmatrix}
0 & 1 \\
-2F'(0) & -2\lambda
\end{pmatrix}, \quad \text{and} \quad
\begin{pmatrix}
0 & 1 \\
-2F'(1) & -2\lambda
\end{pmatrix}.
$$

Since $F'(0) = 1$, the eigenvalues of the Hessian at the fixpoint $(0, 0)$ are $b_\pm = -\lambda \pm \sqrt{\lambda^2 - 2}$, whereas those at the fixpoint $(1, 0)$ are $a_\pm = -\lambda \pm \sqrt{\lambda^2 - 2F'(1)}$. Since of necessity $F'(1) \leq 0$, the eigenvalues $a_\pm$ are both real, but unless $F'(1) = 0$, one is positive and the other negative. Hence $(1, 0)$ is a saddle point. The eigenvalues $b_\pm$ have negative real part, but are real only if $\lambda \geq \sqrt{2}$. In any case, the fix point is stable, but in the case when the eigenvalues are non-real, the solutions of the linearised equations have oscillatory behaviour and cannot stay positive. In the other cases, there exists an integral curve from $(1, 0)$ that approaches $(0, 0)$ along the direction of the smaller of the two eigenvalues, i.e. a map $\gamma : [0, 1] \to \mathbb{R}^2$, such that

$$
\gamma'(\tau) = V(\gamma(\tau)),
$$

where $V(x)$ is in the direction of $W(q, p) \equiv (p, -2f(q) - 2\lambda p)$ but has $|V(x)| \equiv 1$.

In the degenerate case $\lambda = \sqrt{2}$, one can show that the solution analogously has the behaviour of the heavier-tailed solution of the linear equation. From the existence of such an integral curve it follows that for any function $\tau : [0, 1] \to \mathbb{R}$ such that

$$
\tau'(t) = |W(\gamma(\tau(t))))|,
$$

with the property that $\lim_{t \to -\infty} \tau(t) = 0$ and $\lim_{t \to +\infty} \tau(t) = 1$, we have that

$$
w(t) \equiv \gamma(\tau(t)),$
$$

is a solution of $w'(t) = W(w(t))$ and satisfies the right conditions at $\pm \infty$. Clearly the same is true for $\tilde{w}(t) \equiv w(t + a)$, for any $a \in \mathbb{R}$, so solutions are unique only up to a translation. $\square$

We will be mainly interested in the case $\lambda = \sqrt{2}$. The following theorem slightly specialises Bramson’s Theorems A and B from [22] for this case.
Theorem 5.9 ([22]). Let \( u \) be solution of the F-KPP equation (5.4.2) satisfying the standard conditions with \( 0 \leq u(0,x) \leq 1 \). Then there is a function \( m(t) \) such that
\[
 u(t,x + m(t)) \to \omega(x),
\]
uniformly in \( x \), where \( \omega \) is one of the solutions of (22) from Lemma 5.8 as \( t \to \infty \), if and only if
(i) for some \( h > 0 \),
\[
 \limsup_{t \to \infty} \frac{1}{t} \ln \int_t^{t(1+h)} u(0,y)dy \leq -\sqrt{2},
\]
and
(ii) for some \( \nu > 0 \), \( M > 0 \), \( N > 0 \), \( \int_{x-N}^{x+N} u(0,y)dy > \nu \) for all \( x \leq -M \).

Moreover, if \( \lim_{x \to \infty} e^{bx} u(0,x) = 0 \) for some \( b > \sqrt{2} \), then one may choose
\[
 m(t) = \sqrt{2}t - \frac{3}{2\sqrt{2}} \ln t.
\]

Theorem 5.9 is one of the core results of Bramson’s work on Brownian motion and most of the material in his monograph [22] is essentially needed for its proof. We will recapitulate his analysis in the next chapter.

We see that Condition (i) is in particularly satisfied for Heaviside initial conditions. Hence, this result implies that
\[
 P \left( \max_{k \leq m(t)} x_k(t) - m(t) \leq x \right) \to \omega(x) \equiv 1 - w_{\sqrt{2}}(x),
\]
where \( w_{\sqrt{2}} \) solves Eq. (5.5.2).

It should be noted that it follows already from the results of Kolmogorov et al. [54] that (5.5.11) must hold for some function \( m(t) \) with \( m(t)/t \to \sqrt{2} \) (see next chapter). But only Bramson’s precise evaluation of \( m(t) \) shows that the shift is changed from the iid case where it would have to be
\[
 m'(t) = \sqrt{2}t - \frac{1}{2\sqrt{2}} \ln t,
\]
which is larger than \( m(t) \). One can also check that the law of the maximum for the REM on the GW tree, (4.6.7), does not solve (5.5.2). In the next section we will see that this is, however, not so far off.

Remark 5.10. Bramson also shows that convergence to travelling waves takes place if the forward tail of the initial conditions decays more slowly, i.e. if
\[
 \limsup_{t \to \infty} \frac{1}{t} \ln \int_{t}^{t(1+h)} u(0,y)dy = -b,
\]
with \( b < \sqrt{2} \). These cases will not be relevant for us.

Remark 5.11. We will use Bramson’s theorem to proof convergence of Laplace functionals of the extremal process of BBM. Note the for Laplace functionals with
\(\phi\) having bounded support cannot satisfy condition (ii). This precludes uniform convergence, but not pointwise convergence.

### 5.6 The derivative martingale

While Bramson analyses the asymptotic behaviour of the functions \(\omega(x)\), he does not give an explicit form of these solutions, and in particular he does not provide a link to the Gumbel distribution from classical extreme value theory. This was achieved a few years later by Lalley and Sellke (see also the paper by Harris [43]). They wrote a short and very insightful paper [56] that provides, after Bramson’s work, the most significant insights into BBM that we have. It’s main achievement is to give a probabilistic interpretation for the limiting distribution of the maximum of BBM. But this is not all.

For \(1 - f\) satisfying the hypothesis of Theorem 5.9 let \(u\) be given by (5.4.1). Now define

\[
\hat{v}(t, x) \equiv E \left[ \prod_{i=1}^{n(t)} f(\sqrt{2t} + x - x_i(t)) \right] = v(t, \sqrt{2t} + x). \tag{5.6.1}
\]

One checks that \(\hat{v}\) solves the equation

\[
\partial_t \hat{v} = \frac{1}{2} \partial^2_x \hat{v} + \sqrt{2} \partial_x \hat{v} - F(1 - \hat{v}). \tag{5.6.2}
\]

Now let \(\omega\) solve

\[
\frac{1}{2} \partial^2_x \omega + \sqrt{2} \partial_x \omega - F(1 - \omega). \tag{5.6.3}
\]

Then \(\hat{v}(t, x) \equiv \omega(x)\) is a stationary solution to (5.6.2) with initial condition \(\hat{v}(0, x) = \omega(x)\). Therefore we have the stochastic representation

\[
\omega(x) = E \left[ \prod_{i=1}^{n(t)} \omega(\sqrt{2t} + x - x_i(t)) \right]. \tag{5.6.4}
\]

Now probability enters the game.

**Lemma 5.12.** The function \(W(t, x) \equiv \prod_{i=1}^{n(t)} \omega(\sqrt{2t} + x - x_i(t))\) is a martingale with respect to the natural filtration \(\mathcal{F}_t\) of BBM. Moreover, \(W(t, x)\) converges almost surely and in \(L^1\) to a non-trivial limit, \(W^*(x)\), and \(E W^*(x) = \omega(x)\).

**Proof.** The proof is straightforward. It helps to introduce the notation \(x_i^j(t)\) for BBM started in \(y\). Clearly \(W(t, x)\) is integrable. By the Markovian nature of BBM,

\[
W(t+s, x) = \prod_{i=1}^{n(t)} \prod_{j=1}^{n_i(s)} \omega \left( \sqrt{2}(t+s) - x_i^{-x_i(t)}(t) - (x_i^j(s)) \right). \tag{5.6.5}
\]

where the \(x_i^j\) are independent BBM’s. Now
5.6 The derivative martingale

\[ \mathbb{E} [W(t+s,x)|\mathcal{F}_t] = \prod_{i=1}^{n(t)} \mathbb{E} \left[ \prod_{j=1}^{n_i(s)} \omega \left( \sqrt{2}(t+s) - x_j^{-x}(t) - (x^0_{i,j}(s)) \right) \right] |\mathcal{F}_t] \]

\[ = \prod_{i=1}^{n(t)} \mathbb{E} \left[ \prod_{j=1}^{n_i(s)} \omega \left( \sqrt{2}(s) - \left( x_j^{-x}(t) - \sqrt{2}(s) \right) \right) \right] |\mathcal{F}_t] \]

\[ = \prod_{i=1}^{n(t)} \omega \left( \sqrt{2}t - x_j^{-x}(t) \right) = W(t,x). \quad (5.6.6) \]

Now \( W(t,x) \) is bounded and therefore converges almost surely and in \( L^1 \) to a limit, \( W^*(x) \) whose expectation is the \( \omega(x) \).

There are more martingales. First, by a trivial computation,

\[ Y(t) \equiv \sum_{i=1}^{n(t)} e^{\sqrt{2}t - x_i(t)} \quad (5.6.7) \]

is a martingale. Since it is positive and \( \mathbb{E}Y(t) = 1 \), it converges almost surely to a finite non-negative limit. But this means that the exponents in the sum must all go to minus infinity, as otherwise no convergence is possible. This means that

\[ \min_{i \leq n(t)} \left( \sqrt{2}t - x_i(t) \right) \uparrow +\infty, \text{a.s.,} \quad (5.6.8) \]

(this argument is convincing but a bit fishy; see the remark below).

One of Bramson’s results is that (we will see the proof of this in the next chapter, see Corollary 7.2)

\[ 1 - \omega(x) \sim C x e^{-\sqrt{2}x}, \quad \text{as} \quad x \uparrow \infty. \quad (5.6.9) \]

Hence

\[ W(t,x) = \exp \left( \sum_{i=1}^{n(t)} \ln \omega(\sqrt{2}t + x - x_i(t)) \right) \]

\[ \sim \exp \left( -C \sum_{i=1}^{n(t)} (\sqrt{2}t + x - x_i(t)) e^{-\sqrt{2}(\sqrt{2}t + x - x_i(t))} \right) \]

\[ = \exp \left( -C x e^{-\sqrt{2}Y(t)} - C e^{-\sqrt{2}Z(t)} \right), \]

where

\[ Z(t) \equiv \sum_{i=1}^{n(t)} (\sqrt{2}t - x_i(t)) e^{-\sqrt{2}(\sqrt{2}t - x_i(t))}. \quad (5.6.11) \]

\( Z(t) \) is also a martingale, called the derivative martingale, with \( \mathbb{E}Z_t = 0 \). The fact that \( Z(t) \) is a martingale can be verified by explicit computation, but it will actually not be very important for us. In any case, \( Z(t) \) is not even bounded in \( L^1 \) (in fact, an easy calculation shows that \( \mathbb{E}|Z_t| = \frac{2}{\sqrt{\pi}} \) and therefore it is a priori not clear that
$Z(t)$ converges. By the observation (5.6.8), $Z(t)$ is much bigger than $Y(t)$, which implies that unless $Y(t) \to 0$, a.s., it must be true that $\liminf Z(t) = +\infty$, which is impossible since this would imply that $W(t,x) \to 0$, which we know to be false. Hence $Y(t) \to 0$, and thus $Z(t) \to Z$, a.s., where $Z$ is finite and positive. It follows that

$$\lim_{t \to \infty} W(t,x) = \exp \left( -CZe^{-\sqrt{2t}} \right).$$

(5.6.12)

**Remark 5.13.** While the argument of Lalley and Sellke for (5.6.8) may not convince everybody\(^1\), the following gives an alternative proof for the fact that $Y_t \to 0$. By a simple Gaussian computation,

$$\mathbb{P} \left( \exists k \leq n(t) : \sqrt{2t} - x_k(t) < K \right) \leq \frac{e^{\sqrt{2}K}}{\sqrt{4\pi t}}.$$  

(5.6.13)

But this implies that

$$\mathbb{P} \left( \frac{Z(t)}{Y(t)} < K \right) \leq ct^{-1/2}. $$

(5.6.14)

Now assume that $Y_t \to a > 0$, a.s.. Then, for large enough $t$,

$$\mathbb{P} \left( Z(t) < 2K a \right) \leq ct^{-1/2}.$$  

(5.6.15)

But this implies that

$$\mathbb{P} \left( \liminf_n Z(2^n) < 2aK \right) = 0.$$  

(5.6.16)

and hence

$$\mathbb{P} \left( \limsup_{t \to \infty} Z(t) < 2aK \right) = 0.$$  

(5.6.17)

But this implies that

$$\liminf_{t \to \infty} W(t,x) = 0, a.s.,$$

(5.6.18)

and since we know that $W(t,x)$ converges almost surely, it must hold that the limit is zero. But this is in contradiction with the fact that the limit is a non-negative random variable with positive expectation and that convergence holds in $L^1$. Hence it must be the case that $Y_t \to 0$, almost surely.

**Remark 5.14.** One may interpret $Y(t)$ as a partition function. Namely, if we set

$$Z_\beta(t) = \sum_{i=1}^{n(t)} e^{\beta x_i(t)},$$

(5.6.19)

then

$$Y(t) = \frac{Z_{\sqrt{2}}(t)}{\mathbb{E} Z_{\sqrt{2}}(t)}.$$  

(5.6.20)

---

\(^1\) Thanks to Marek Biskup for voicing some doubt about this claim.
\( \sqrt{2} \) has the natural interpretation of the **critical inverse temperature** for this model, which can be interpreted as the value where the "law of large numbers" starts to fail in a strong sense, namely that \( Z_\beta(t) / \mathbb{E}Z_\beta(t) \) does no longer converge to a non-trivial limit. In the REM, the critical value is \( \sqrt{2 \ln 2} \), and it was shown in \( \cite{21} \) that in this case, at the critical value, this ratio converges to \( 1/2 \). For BBM, one can show that

\[
\frac{Z_\beta(t)}{\mathbb{E}Z_\beta(t)},
\]  
(5.6.21)

is a uniformly integrable martingale for all values \( \beta < 1 \) that converges a.s. and in \( L^1 \) to a positive random variable of mean 1. The reason for the name **derivative martingale** is that \( Z_t \) looks like the derivative of the martingale \( Y_t \) with respect to \( \beta \) at the value \( \beta = \sqrt{2} \). This is indeed a strange animal: its limit is almost surely positive, but the limit of its expectation is zero.

**Remark 5.15.** Convergence of the derivative martingale is shown via purely probabilistic techniques by Kyprianou \( \cite{55} \).

Finally, we return to the law of the maximum. We have that, for any \( s \geq 0 \),

\[
\lim_{t \uparrow \infty} \mathbb{P} \left( \max_{k \leq n(t)} x_k(t) - m(t) \leq x \mid \mathcal{F}_s \right) = \lim_{t \uparrow \infty} \mathbb{P} \left( \max_{k \leq m(t)+s} x_k(t+s) - m(t+s) \leq x \mid \mathcal{F}_s \right) = \lim_{t \uparrow \infty} \mathbb{P} \left( \max_{k \leq m(t)+s} x_k(t+s) - m(t+s) \leq x \mid \mathcal{F}_s \right)
\]

\[
= \lim_{t \uparrow \infty} \prod_{i=1}^{m(s)} \mathbb{P} \left( \max_{k \leq m(t)} x_{i,k(t)} - m(t+s) - x_i(s) \leq x \mid \mathcal{F}_s \right) = \lim_{t \uparrow \infty} \prod_{i=1}^{m(s)} u(t, x + m(t+s) - x_i(s)).
\]  
(5.6.22)

where \( 1 - u \) is the solution of the F-KPP equation with Heaviside initial condition. Next we use that

\[ m(t+s) - m(t) - \sqrt{2} s \to 0, \quad \text{as} \ t \uparrow \infty, \]  
(5.6.23)

for fixed \( s \). This shows that

\[
\lim_{t \uparrow \infty} \mathbb{P} \left( \max_{k \leq n(t)} x_k(t) - m(t) \leq x \mid \mathcal{F}_s \right) = \lim_{t \uparrow \infty} \prod_{i=1}^{m(s)} u(t, x + m(t) - \sqrt{2} s - x_i(s)) = \prod_{i=1}^{m(s)} \omega(x - \sqrt{2} s - x_i(s)) = W(s, x).
\]
As now \( s \uparrow \infty \),

\[
W(s,x) \to e^{-Cze^{-\sqrt{2}x}}, \text{ a.s.} \quad (5.6.24)
\]

which proves the main theorem of Lalley and Sellke [56]:

**Theorem 5.16 ([56]).** For BBM,

\[
\lim_{t \uparrow \infty} \mathbb{P} \left( \max_{k \leq n(t)} x_k(t) - m(t) \leq x \right) = e^{-Cze^{-\sqrt{2}x}}. \quad (5.6.25)
\]

Moreover

\[
\lim_{s \uparrow \infty} \lim_{t \uparrow \infty} \mathbb{P} \left( \max_{k \leq n(t)} x_k(t) - m(t) \leq x | \mathcal{F}_s \right) = e^{-Cze^{-\sqrt{2}x}}, \text{ a.s.} \quad (5.6.26)
\]

**Remark 5.17.** Of course, the argument above shows that any solution of (5.5.2) satisfying the conditions of Lemma 5.8 has a representation of the form \( 1 - \mathbb{E} \left( e^{-Cze^{-\sqrt{2}x}} \right) \), with only different constants \( C \).

**Remark 5.18.** Lalley and Sellke conjectured in [56] that

\[
\lim_{T \uparrow \infty} \frac{1}{T} \int_0^T \mathbb{1}_{\max_{k \leq n(t)} x_k(t) - m(t) \leq x} dt \to e^{-Cze^{-\sqrt{2}x}}, \text{ a.s.} \quad (5.6.27)
\]

This was proven to be true in [5].
Chapter 6
Bramson’s analysis of the F-KPP equation

In this chapter we recapitulate the essential features of Bramson’s proof of his main theorems on travelling wave solutions. The material is taken from [22] with occasionally a few details added in proofs and some omissions.

6.1 Feynman-Kac representation

Bramson’s analysis of the asymptotics of solutions of the KPP equation in [22] relies on a Feynman-Kac representation. The derivation of this can be seen as a combination of the the mild formulation of the pde together with the standard Feynman-Kac representation of the heat kernel. Namely, if $u(t,x)$ is the solution of the linear equation

$$\partial_t u = \frac{1}{2}\partial_x^2 u + k(t,x)u,$$

with initial conditions $u(0,x)$, then the standard Feynman-Kac formula yields the representation ([47], see e.g. [70, 51])

$$u(t,x) = \mathbb{E}_x \left( \exp \left( \int_0^t k(t-s, B_s) ds \right) u(0, B_t) \right),$$

where the expectation is with respect to ordinary Brownian motion started at $x$. To get back to the full equation, we use this with

$$k(t,x) = F(u(t,x))/u(t,x),$$

where $u$ itself is the solution of the F-KPP equation. It may seem that this is just a re-writing of the original equation as an integral equation. Still, the ensuing representation is very useful as it allows to process a priori information on the solution into finer estimates. Note that under the \textit{standard conditions on $F$} stated in (5.4.10) and (5.4.11),

$$0 \leq k(t,x) \leq 1$$
and
\[ k(t, x) \sim Cu(t, x)^p, \quad \text{when } u(t, x) \downarrow 0. \quad (6.1.5) \]

Bramson’s key idea was to use the FK-representation not all the way down to the initial condition, but to go back to some arbitrary time \( 0 \leq r < t \).

**Theorem 6.1.** Let \( u \) be the solution of the F-KPP equation with initial condition \( u(0, x) \). Let \( k(t, x) \equiv F(u(t, x))/u(t, x) \). Then, for any \( 0 \leq r \leq t \), \( u \) satisfies the equation
\[ u(t, x) = \mathbb{E}_x \left( \exp \left( \int_0^{t-r} k(t-s, B_s) ds \right) u(r, B_{t-r}) \right), \quad (6.1.6) \]
where \( (B_s, s \in \mathbb{R}_+) \) is Brownian motion starting in \( x \).

This can be conveniently be rewritten in terms of Brownian bridges. Recall the definition of a Brownian bridge \((s^x_{t, y}, 0 \leq s \leq t)\) from \( x \) to \( y \) in time \( t \). Then (6.1.6) can be written as
\[ u(t, x) = \int_{-\infty}^{\infty} dy \mathbb{E}_y \left( \exp \left( \int_0^{t-r} k(t-s, s^x_{t, y}(s)) ds \right) \right), \quad (6.1.7) \]

Note that under the standard conditions \( k(t, x) \) lies between zero and one and tends to one as \( \infty \). Bramson’s basic idea to exploit this formula is to prove a priori estimates on the function \( k(t, x) \). Note that \( k(t, x) > 0 \), and by Condition 5.4.11, \( k(x, t) \sim 1 \), when \( u(t, x) \sim 0 \) (think about the simplest case when \( k(t, x) = 1 - u(t, x) \)). We will see that \( k(s, x) \sim 1 \) if \( x \) is a above a certain curve, \( \mathcal{M} \). On the other hand, he shows that the probability that the Brownian bridge descends below a curve \( \mathcal{M} \) is negligibly small. The following proposition is this strategy and essentially contained in Bramson’s monograph [22, Proposition 8.3].

**Theorem 6.2.** Let \( u \) be a solution to the F-KPP equation (5.4.2) with initial condition satisfying
\[ \int_0^\infty y e^{\sqrt{2}y} u(0, y) dy < \infty. \quad (6.1.8) \]

Define
\[ \psi(r, t, z + \sqrt{2}t) = \frac{e^{-\sqrt{2}t}}{\sqrt{2\pi(t-r)}} \int_0^\infty u(r, y + \sqrt{2}r) e^{\sqrt{2}y} e^{-\frac{(y - \frac{t-r}{2\sqrt{2}})^2}{2\frac{t-r}{2\sqrt{2}}}} \left( 1 - e^{-2y} \right) \right) dy \quad (6.1.9) \]

Then for \( r \) large enough, \( t \geq 8r \), and \( z \geq 8r - \frac{3}{2\sqrt{2}} \ln t \),
\[ \gamma^{-1}(r) \psi(r, t, z + \sqrt{2}t) \leq u(t, z + \sqrt{2}t) \leq \gamma(r) \psi(r, t, z + \sqrt{2}t), \quad (6.1.10) \]
where \( \gamma(r) \downarrow 1, \text{ as } r \to \infty. \)

---

1 One should appreciate the beauty of this construction: start with a probabilistic model (BBM), derive a pde whose solutions represent quantities of interest, and then use a different probabilistic representation of the solution (in terms of Brownian motion) to analyse these solutions...
We see that $\psi$ basically controls $u$, but of course $\psi$ still involves $u$. The fact that $u$ is bounded by $u$ may seems strange, but we shall see that this is very useful.

**Proof.** The better part of this chapter will be devoted to proving the following two facts:

1. For $r$ large enough, $t \geq 8r$ and $x \geq m(t) + 8r$

\[
u(t,x) \geq \nu_1(r,t,x)
\]

\[\equiv C_1(r) e^{-r} \int_{-\infty}^{\infty} u(r,y) \frac{e^{-(x-y)^2/2\pi(r-r)}}{\sqrt{2\pi(r-r)}} \mathbb{P}\left(\xi_{t,r}^{x-y}(s) > \mathcal{M}_{t,r}^{x}(t-s), s \in [0,t-r]\right) dy
\]

and

\[
u(t,x) \leq \nu_2(r,t,x)
\]

\[\equiv C_2(r) e^{-r} \int_{-\infty}^{\infty} u(r,y) \frac{e^{-(x-y)^2/2\pi(r-r)}}{\sqrt{2\pi(r-r)}} \mathbb{P}\left(\xi_{t,r}^{x-y}(s) > \mathcal{M}_{t,r}^{x}(t-s), s \in [0,t-r]\right) dy,
\]

where the functions $\mathcal{M}_{t,r}^{x}(t-s), \mathcal{M}_{t,r}^{x}(t-s)$ satisfy

\[\mathcal{M}_{t,r}^{x}(t-s) \leq n_{r,t}(t-s) \leq \mathcal{M}_{t,r}^{x}(t-s).
\]

Here

\[n_{r}(s) \equiv \sqrt{2r} + \frac{(s-r)}{t-r} (m(t) - \sqrt{2r}).
\]

Moreover, $C_1(r) \uparrow 1, C_2(r) \downarrow 1$ as $r \uparrow \infty$.

2. The bounds $\nu_1(r,t,x)$ and $\nu_2(r,t,x)$ satisfy

\[1 \leq \frac{\nu_2(r,t,x)}{\nu_1(r,t,x)} \leq \gamma(r),
\]

where $\gamma(r) \downarrow 1$ as $r \uparrow \infty$.

Assuming these facts, we can conclude the proof of the theorem rather quickly. Set

\[
u_1(t,x) = e^{-r} \int_{-\infty}^{\infty} u(r,y) \frac{e^{-(x-y)^2/2\pi(r-r)}}{\sqrt{2\pi(r-r)}} \mathbb{P}\left(\xi_{t,r}^{x-y}(s) > n_{r,t}(t-s), s \in [0,t-r]\right) dy
\]

By (6.1.13), we have $\nu_1 \leq \nu \leq \nu_2$. Therefore, for $r,t$ and $x$ large enough

\[rac{u(t,x)}{\nu_1(t,x)} \leq \frac{\nu_2(r,t,x)}{\nu(r,t,x)} \leq \frac{\nu_2(r,t,x)}{\nu_1(r,t,x)} \leq \gamma(r),
\]

and
Combining (6.1.17) and (6.1.18) we get

\[ \gamma^{-1}(r)\psi(r,t,x) \leq u(t,x) \leq \gamma(r)\psi(r,t,x). \]  

(6.1.19)

For \( x \geq 8r - \frac{1}{2}\sqrt{2} \ln t \) we get from (6.1.19) that

\[ \gamma^{-1}(r)\psi(r,t,x + \sqrt{2}t) \leq u(t,x + \sqrt{2}t) \leq \gamma(r)\psi(r,t,x + \sqrt{2}t). \]  

(6.1.20)

The nice thing is that the probability involving the Brownian bridge in the definition of \( \psi \) can be explicitly computed, see Lemma 5.1.

Since our bridge goes from \( x + \sqrt{2}t \) to \( y \) and the straight line is between \( \sqrt{2}r \) and \( m(t) \), and it stays above this line, we have to adjust this to length \( t - r \), \( y_1 = \sqrt{2}t + x - m(t) = x + \frac{1}{2}\sqrt{2} \ln t > 0 \) for \( t > 1 \) and \( y_2 = y - \sqrt{2}r \).

Of course \( \mathbb{P}\left( \xi_{\gamma^{-1}}(s) > n_{t,s}(t-s), s \in [0,t-r) \right) = 0 \) for \( y \leq \sqrt{2}r \). For \( y > \sqrt{2}r \), from Lemma 5.1 we have

\[ \mathbb{P}\left( \xi_{\gamma^{-1}}(s) > n_{t,s}(t-s), s \in [0,t-r) \right) = 1 - \exp\left( -\left( x + \frac{3}{2\sqrt{2}} \ln t \right) \left( y - \sqrt{2}r \right) \right). \]  

(6.1.21)

Changing variables to of \( y' = y - \sqrt{2}r \) in the integral appearing in the definition of \( \psi \), we get (6.1.9). This, together with (6.1.19), concludes the proof of the proposition. \( \square \)

**Remark 6.3.** In a later paper [23], Bramson gives an even more explicit representation for the asymptotics of the solutions with speed \( \sqrt{2} \), namely,

\[ u(t,z+m(t)) = 2C(t,z)\gamma^{-1} e^t \int_0^\infty yu(0,y) \frac{1}{\sqrt{2\pi t}} \exp \left( -(z+m(t)-y)^2/2t \right) dy, \]  

(6.1.22)

with \( \lim_{t \to \infty} \lim_{r \to \infty} C(t,z) = 1 \). We will not make use of this and therefore also do not give the proof.

In the remainder of this chapter we will recall Bramson’s analysis of the F-KPP equation that leads to the proof of the facts we used in the proof above. In the subsequent chapter we use this proposition crucially to obtain the extremal process of BBM. Let us make some comments about what needs to be shown. In the bounds (6.1.11) and (6.1.12), the exponential factor involving \( k(s,x) \) in the Feynman-Kac representation (6.1.7) is replaced by 1. For an upper bound, this is easy, since \( k \leq 1 \); it is, however, important that we are allowed to smuggle in the condition that the Brownian bridge stays above the curve \( \mathcal{M}_{t,s} \), which is lying somewhat, but not too much, below the straight line \( n_{t,s} \). For the lower bound, we can of course introduce a condition for the Brownian bridge to stay above the curve \( \mathcal{M}_{t,s} \), which lies above the straight line. We then have to show that then \( k(t-s,\xi_{\gamma^{-1}}(s)) \) is very close to
6.2 The maximum principle and its applications

1. For this we seem to need to know how the solution $u(t,x)$ behaves, but, in fact, an upper bound suffices. But a trivial upper bound is always given by the solution of the linear F-KPP equation for which the Feynman-Kac representation is explicit and allows to get very precise bounds. Finally, one needs to show that the two probabilities concerning the Brownian bridge are almost the same. This holds because of what is known as entropic repulsion: A Brownian bridge wants to fluctuate on the order $\sqrt{t}$. If one imposes a condition not to make such fluctuations in the negative direction, the bridge will prefer to stay positive and in fact behaves as if one would force it to stay above a curve that goes only up to some $t^\delta$, with $\delta < 1/2$. The remainder of this chapter explains how these simple ideas are implemented.

6.2 The maximum principle and its applications

Let us first remark that under the standard conditions on $F$, existence and uniqueness of solutions of the $F-KPP$ equation follow easily via the standard tools of Picard iteration and the Gronwall lemma.

**Theorem 6.4.** If $F$ satisfies the standard conditions, then the $F-KPP$ equation with initial data $0 \leq u_0(x) \leq 1$ has a unique solution for all $t \in \mathbb{R}_+$ and $0 \leq u(t,x) \leq 1$, for all $t \in \mathbb{R}_+$ and $x \in \mathbb{R}$.

**Remark 6.5.** Existence and uniqueness hold for general bounded initial conditions, if $F$ is assumed to be Lipschitz, but we are only interested in solutions that stay in $[0,1]$.

**Proof.** The best way to prove existence is to set up a Picard iteration for the mild form of the equation, i.e. to define $u^1(t,x)$ as the unique solution of the heat equation

$$\partial_t u^1(t,x) = \frac{1}{2} \partial^2_{xx} u(t,x), \quad t \in \mathbb{R}_+, x \in \mathbb{R}, \quad u^1(0,x) = u_0(x).$$

(6.2.1)

This can be written as

$$u^1(t,x) = \int_0^t \int_{-\infty}^{\infty} g(t-s,x-y)u_0(y),$$

(6.2.2)

where

$$g(t,x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}},$$

(6.2.3)

is the heat kernel. Then define recursively

$$u^n(t,x) = \int_0^t ds \int_{-\infty}^{\infty} g(t-s,x-y)u_0(y) + \int_0^t ds \int_{-\infty}^{\infty} dy g(t-s,x-y)F(u^{n-1}(s,y)) \equiv T(u^{n-1})(t,x).$$

(6.2.4)

If this sequence converges, the limit will be a solution of
which is a mild formulation of the F-KPP equation. To prove convergence, one shows that the map \( T \) is a contraction in the sup-norm (on \([0,t_0] \times \mathbb{R}\), for \( t_0 \) sufficiently small. This follows easily since \( F \) is Lipschitz continuous, so that

\[
|T(u)(t,x) - T(v)(t,x)| \leq \int_0^t ds \int_{-\infty}^{\infty} dy g(t-s,x-y) C|u(s,y) - v(s,y)| \quad (6.2.6)
\]

Hence,

\[
\sup_{x \in \mathbb{R}, t \in [0,t_0]} |T(u)(t,x) - T(v)(t,x)| \leq C t_0 \sup_{y \in \mathbb{R}, s \in [0,t_0]} |u(s,y) - v(s,y)|, \quad (6.2.7)
\]

by which \( T \) is a contraction of \( t_0 < 1/C \). This proves that a solution exists up to time \( 1/C \). From here we may start with \( u(0,x) \) as a new initial contain to construct the solution up to time \( 2/C \), and so on. This shows global existence.

To prove uniqueness, we could use the Gronwall lemma, but it follows immediately from the maximum principle which we state next.

One of the fundamental tools in Bramson’s analysis is the following maximum principle for solutions of the F-KPP equation. We take the theorem from Bramson who attributes it to Aronson and Weinberger [7].

**Proposition 6.6.** Let \( F \) satisfy the standard conditions (5.4.10). Assume that \( u^1(t,x) \) and \( u^2(t,x) \) satisfy the inequalities

\[
0 \leq u^1(0,x) \leq u^2(0,x) \leq 1, \quad \forall x \in (a,b), \quad (6.2.8)
\]

and

\[
\partial_t u^2(t,x) - \frac{1}{2} \partial_x^2 u^2(t,x) - F(u^2(t,x)) \geq \partial_t u^1(t,x) - \frac{1}{2} \partial_x^2 u^1(t,x) - F(u^1(t,x)), \quad (6.2.9)
\]

for all \((t,x) \in (0,T) \times (a,b)\). If \( a > -\infty \), assume further that

\[
0 \leq u^1(t,a) \leq u^2(t,a) \leq 1, \quad \forall t \in [0,T], \quad (6.2.10)
\]

and if \( b < \infty \), assume that

\[
0 \leq u^1(t,b) \leq u^2(t,b) \leq 1, \quad \forall t \in [0,T]. \quad (6.2.11)
\]

Then \( u^2 \geq u^1 \), and if the inequality (6.6) is strict in an open interval of \((a,b)\), then \( u^2 > u^1 \) on \((0,T) \times (a,b)\).
Proof. The proof consists of reducing the argument to the usual maximum principle for parabolic linear equations. Condition (6.2.9) can be written as

$$\partial_t(u^2 - u^1) - \frac{1}{2} \partial_x^2 (u^2 - u^1) \geq F(u^2) - F(u^1) = F(u^1 + \theta(u^2 - u^1))(u^2 - u^1),$$

(6.2.12)

for some $\theta \in [0,1]$, by the mean value theorem. Now set $\alpha \equiv \max_{u \in [0,1]} F'(u)$, and set

$$v(t,x) \equiv e^{-2\alpha t} (u^2(t,x) - u^1(t,x)).$$

(6.2.13)

Then by (6.2.12), $v$ satisfies the differential inequality

$$\partial_t v(t,x) - \frac{1}{2} \partial_x^2 v(t,x) \geq (F'(u^1 + \theta(u^2 - u^1)) - 2\alpha) v(t,x),$$

(6.2.14)

where by construction the coefficient to $v$ on the right hand side is non-positive. If now for some $(t_0,x_0)$, $v(t_0,x_0) < 0$, then there must be some point $(t_1,x_1)$, with $t_1 \leq t_0$, such that $v$ takes its minimum in $[0,t_0] \times (a,b)$ at this point. Note that $x_1$ cannot be in the boundary of $(a,b)$, due to the boundary conditions. But then at $(t_1,x_1)$, $\partial_t v \leq 0$ and $\partial_x^2 v \geq 0$, so that the left hand side of (6.2.12) is non-positive, whereas the left hand side is strictly positive, which leads to a contradiction. Hence $v \geq 0$, and so $u^2 \geq u^1$. \(\Box\)

Clearly, the preceding theorem implies that two solutions of the F-KPP equation with ordered initial (and boundary) conditions remain ordered for all times. Since, moreover, the constant functions 0 and 1 are solutions of the F-KPP equations, any solution with initial conditions within [0,1] will remain in this interval. This follows, of course, also from the McKean representation (5.4.1).

The following corollary compares solutions of the linear and non-linear equation.

**Corollary 6.7 (Corollary 1 of [22]).** Let $u^1$ be a solution of the F-KPP equation with $F$ satisfying (5.4.10) and let $u^2$ satisfy the linear equation

$$\partial_t u^2(t,x) = \frac{1}{2} \partial_x^2 u^2(t,x) + u^2(t,x).$$

(6.2.15)

If $u^1(0,x) = u^2(0,x)$ for all $x$, then

$$u^1(t,x) \leq u^2(t,x),$$

(6.2.16)

for all $t > 0, x \in \mathbb{R}$. Therefore, if $u^1(0,x) = 1_{x \leq 0}$, the re-centring $m(t)$ satisfies

$$m(t) < \sqrt{2t} - 2^{-3/2} \ln t + C = \tilde{m}(t) + C.$$

(6.2.17)

**Proof.** Since $F(u) \leq u$, the two functions satisfy the hypothesis of Proposition 6.6. Thus, (6.2.16) holds. On the other hand, one can solve $u^2$ explicitly, using the heat kernel, as

$$u^2(t,x) = \frac{e^t}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-(x-y)^2/(2t)} u^2(0,y) dy.$$
For Heaviside initial condition this becomes
\[
    u^2(t,x) = \frac{e^t}{\sqrt{2\pi t}} \int_{-\infty}^{0} e^{-(x-y)^2/(2t)} dy = e^t \mathbb{P}(N(0,t) > x). \tag{6.2.19}
\]

Thus, we know that
\[
    u^2(t,x + \tilde{m}(t)) \to e^t e^{-\epsilon_2}. \tag{6.2.20}
\]

By monotonicity, we see that \(m(t)\) must be less than \(\tilde{m}(t)\), up to constants. Of course, we know this probabilistically already.

An extension of the maximum principle due to McKean also controls the sign changes of solutions.

**Proposition 6.8 (Proposition 3.2 in [22])**. Let \(u^i, i = 1,2\) be solutions of the F-KPP equation with \(F\) satisfying (5.4.10). Assume that the initial conditions are such that, if \(x_1 < x_2\), then \(u^2(0,x_1) > u^2(0,x_2)\) implies that \(u^2(0,x_2) \geq u^1(0,x_1)\). Then this property is preserved for all times, i.e. for all \(t > 0\), if \(x_1 < x_2\), then \(u^2(t,x_1) > u^2(t,x_2) \geq u^1(t,x_1) \geq u^1(t,x_2)\).

**Proof.** We set \(v = u^2 - u^1\). Then
\[
    \partial_t v = \frac{1}{2} \partial_x^2 v + F'(u^1 + \theta(u^2 - u^1))v. \tag{6.2.21}
\]

We now use the Feynman-Kac representation in the form (6.1.6) with
\[
    k(t,x) \equiv F'(u^1(t,x) + \theta(t,x)(u^2(t,x) - u^1(t,x))), \tag{6.2.22}
\]

the coefficient of \(v\) on the right of (6.2.21). Set
\[
    M(r) \equiv \exp \left( \int_0^r k(t-s,B_x)ds \right) v(t-r,B_r). \tag{6.2.23}
\]

Then \(v(t,x) = \mathbb{E}_x[M(r)]\), for any \(t \geq r \geq 0\); in fact, \(M(r)\) is a martingale. Next we replace the constant time \(r\) by the stopping time
\[
    \tau \equiv \inf\{0 \leq s \leq t : M(s) = 0\} \wedge t. \tag{6.2.24}
\]

Now choose an \(x_1\) such that \(v(t,x_1) > 0\). Then
\[
    0 < v(t,x_1) = \mathbb{E}_{x_1}[M(\tau)], \tag{6.2.25}
\]

so that \(\mathbb{P}(M(\tau) > 0) > 0\). But this means that, with positive probability \(\tau = t\), which can only happen if there is a continuous path, \(\gamma\), starting at \(x_1\), such that for all \(0 \leq s \leq t\), \(v(t-s,\gamma(s)) > 0\). In particular, \(v(0,\gamma(t)) > 0\). Now define another stopping time, \(\sigma\), by
\[
    \sigma \equiv \inf\{0 \leq s \leq t : B_x = \gamma(s)\} \wedge t, \tag{6.2.26}
\]
where \( B_0 = x_2 \). Now \( M(\sigma) \geq 0 \). If \( \sigma < t \), the inequality is strict, and if \( \sigma = t \), it must be the case that \( B(t) \geq \gamma(t) \), since otherwise the BM would have had to cross \( \gamma \). But then \( \nu(0,B(t)) > 0 \) by assumption on the initial conditions. On the other hand, we have again,

\[
v(t,x_2) = \mathcal{E}_{x_2}[M(\sigma)].
\] (6.2.27)

But Brownian motion starting at \( x_2 \) hits the path \( \gamma \) with positive probability before time \( t \), so the left-hand side of (6.2.27) is strictly positive, hence \( v(t,x_2) > 0 \).

The first main goal is to prove an a priori convergence result to travelling wave solutions that was already known to Kolmogorov, Petrovsky, and Piscounov [54]. We will use the notation

\[
m_\varepsilon(t) \equiv \sup\{x \in \mathbb{R} : u(t,x) \geq \varepsilon\}.
\] (6.2.28)

**Theorem 6.9** ([54]). Let \( u \) be a solution of the F-KPP equation with \( F \) satisfying (5.4.10) and Heaviside initial conditions. Then

\[
u(t,x + m_{1/2}(t)) \to w^{\sqrt{2}}(x),
\] (6.2.29)

uniformly in \( x \), as \( t \uparrow \infty \), where \( w^{\sqrt{2}} \) is the solution of the stationary KPP equation

\[
\frac{1}{2}w'' + \sqrt{2}w' + F(w) = 0,
\] (6.2.30)

with \( w(0) = 1/2 \). Moreover,

\[
m_{1/2}(t)/t \to \sqrt{2}.
\] (6.2.31)

**Proof.** We need two results that are consequences of Proposition 6.8.

**Corollary 6.10** (Corollary 1 in [22]). For \( u \) as in Theorem 6.9, for any \( 0 < \varepsilon < 1 \),

\[
\begin{align*}
&u(t,m_\varepsilon(t) + x) \uparrow w(x), & \text{for } x > 0, \\
&u(t,m_\varepsilon(t) + x) \downarrow w(x), & \text{for } x \leq 0,
\end{align*}
\] (6.2.32)

for some functions \( w \).

**Proof.** We fix \( t_0, a \in \mathbb{R}_+ \) and set

\[
\begin{align*}
u^1(t,x) & \equiv u(t,x + m_\varepsilon(t_0)), \\
u^2(t,x) & \equiv u(t + a,x + m_\varepsilon(t_0 + a)).
\end{align*}
\] (6.2.33)

Clearly these functions satisfy the hypothesis of Proposition 6.8. Moreover, for \( u(t,m_\varepsilon(t)) = \varepsilon \), and so

\[
u^2(t_0,0) = u^1(t_0,0).
\] (6.2.34)

Hence, by Proposition 6.8, for \( x \geq 0 \), \( u^2(t_0,x) \geq u^1(t_0,x) \), i.e.

\[
u(t_0 + a,x + m_\varepsilon(t_0 + a)) \geq u(t_0,x + m_\varepsilon(t_0)).
\] (6.2.35)
Likewise, for \( x < 0 \), if for some \( x < 0 \), it were true that \( u^2(t_0,x) > u^1(t_0,x) \), then the same would need to hold at 0, which is not the case. Hence, \( u^2(t_0,x) \leq u^1(t_0,x) \), i.e.

\[
u(t_0 + a, x + m(x)(t_0 + a)) \leq u(t_0, x + m(x)(t_0)).
\]

(6.2.36)

Hence \( u(t, m(x)(t) + x) \) is monotone increasing and bounded from above for \( x > 0 \), and decreasing and bounded from below for \( x \leq 0 \). This implies that it must converge to some function \( w \), as claimed. \( \Box \)

We need one more corollary that involves the notion of being more stretched. For \( g, h \) monotone decreasing functions, \( g \) is more stretched than \( h \), if for any \( c \in \mathbb{R} \) and \( x_1 < x_2, g(x_1) > h(x_1 + c) \) implies \( g(x_2) > h(x_2 + c) \), and the same with \( > \) replaced by \( \geq \).

**Corollary 6.11 (Corollary 2 in [22]).** If \( u' \) are solutions of the F-KPP equations as in the Theorem, then if the initial conditions of \( u^2 \) are more stretched than those of \( u^1 \), then for any time, \( u^2 \) is more stretched than \( u^1 \). In particular, if \( u^1 \) has Heaviside initial conditions, than \( w^\sqrt{2} \) is more stretched than \( u^1(t, \cdot) \) for all times.

**Proof.** Straightforward from Proposition 6.8 \( \Box \)

So now we know that \( u(t, m_{1/2}(t) + x) \) converges to some function, \( w(x) \). moreover, since \( u \) is monotone, so is the limit. Moreover, \( u(t, x) \) is less stretched than \( w \). Thus \( w(x) \leq w^\sqrt{2}(x) \), for \( x > 0 \), and \( w(x) \geq w^\sqrt{2}(x) \), for \( x \leq 0 \). It is not hard to show that the limit is indeed \( w^\sqrt{2} \); if we set \( v(t, x) \equiv u(t, x + m(t)) \), then

\[
\partial_t v(t, x) = \partial_t u(t, m(t) + x) + m'(t) \partial_x u(t, m(t) + x) = \frac{1}{2} \partial_x^2 v + m'(t) \partial_x v + F(v).
\]

(6.2.37)

Integrating the left-hand side over \( x \) twice starting from zero, and then integrating the result over \( t \) over an interval of length 1, we get

\[
\int_0^x dy \int_0^y dz (u(t + 1, z + m(t + 1)) - u(t, z + m(t))) \leq \frac{x^2}{2} \sup_{0 \leq s \leq t} (u(t + 1, z + m(t + 1)) - u(t, z + m(t))),
\]

(6.2.38)

which tends to zero by Corollary 6.10, as \( t \uparrow \infty \). Applying the same procedure on the right-hand side yields

\[
\int_0^t ds \int_0^x dy \int_0^y dz \left( \frac{1}{2} \partial_x^2 v + m'(t) \partial_x v + F(v) \right)
\]

(6.2.39)

\[
= \int_0^t ds \left( \frac{1}{2} v(s, x) - \frac{1}{2} v(s, 0) - \frac{x}{2} \partial_x v(s, 0) + m'(s) \int_0^x dy (v(s, y) - v(s, 0)) + \int_0^y dz F(v(s, z)) \right).
\]

Using the convergence of \( v \) to \( w \), as \( t \uparrow \infty \), and setting \( a \equiv \lim_{t \rightarrow \infty} \partial_x v(t, x) \), the first line converges to
\[
\int_t^{t+1} ds \left( \frac{1}{2} v(s,x) - \frac{1}{2} v(s,0) - \frac{x}{2} \partial_x v(s,0) \right) \to \frac{1}{2} w(x) - \frac{1}{4} - \frac{a}{2} x.
\] (6.2.40)

Next, for the same reason,
\[
\int_0^x dy (v(s,y) - v(s,0)) \to \int_0^x \left( w(y) - \frac{1}{2} \right) dy.
\] (6.2.41)

Finally,
\[
\int_t^{t+1} ds \int_0^x dy \int_0^y dz F(v(s,z)) \to \int_0^x dy \int_0^y dz F(w(z)).
\] (6.2.42)

Since all of (6.2.39) tends to zero, also the term involving \( m'(s) \) must be independent of \( t \), asymptotically. This means that
\[
\int_t^{t+1} ds m'(s) = m(t+1) - m(t) \to \lambda,
\] (6.2.43)

for some \( \lambda \in \mathbb{R} \). Putting all this together we get that
\[
\frac{1}{2} w(x) - \frac{1}{4} - \frac{a}{2} x + \lambda \int_0^x \left( w(y) - \frac{1}{2} \right) dy + \int_0^x dy \int_0^y dz F(w(z)) = 0.
\] (6.2.44)

Differentiating twice with respect to \( x \) yields that \( w \) satisfies the travelling wave equation (5.5.2) for some value of \( \lambda \geq \sqrt{2} \). But we have already an upper bound on \( m \) which implies that \( \lambda \leq \sqrt{2} \), which leaves us with \( \lambda = \sqrt{2} \). This concludes the proof of the theorem. \( \square \)

The maximum principle has several important consequences for the asymptotics of solutions that will be exploited later. The first is a certain monotonicity property.

**Lemma 6.12** (Lemma 3.1 in [22]). Let \( u \) be a solution of the F-KPP equation as before. Assume that the initial conditions satisfy \( u(0,x) = 0 \), for \( x \geq M \), for some real \( M \). Then, for any \( t \), for all \( y \geq x \geq M \),
\[
u(t,x) \leq \nu(t,2x - y),
\] (6.2.45)

and therefore,
\[
\partial_x u(t,x) \leq 0.
\] (6.2.46)

**Proof.** Let \( u^i, i = 1, 2, \) be solutions on \( \mathbb{R}_+ \times [x, \infty) \) with initial data \( u^1(0,y) = 0 \), for \( y \geq x \), and \( u^2(0,y) = u(0,2x - y) \) for \( y \geq x \), and with boundary conditions \( u^i(t,x) = u(t,x), \) for all \( t \geq 0 \). By the maximum principle, \( u^1 \leq u^2 \), for all times. Since \( x \geq M \), and hence \( u^1(t,y) = u(t,y) \), for \( y \geq x \). On the other hand, by reflection symmetry, \( u^2(y,y) = u(y,2x - y) \). Hence (6.2.45). Since this inequality implies that \( u(t,x + \varepsilon) \leq u(t,x - \varepsilon) \), the claim on the derivative at \( x \) is immediate. \( \square \)

The next lemma gives a continuity result for the solutions as functions of the initial conditions.
Lemma 6.13 (Lemma 3.2 in [22]). Let \( u^i \) be solutions of the F-KPP equation with \( F'(u) \leq 1 \), and assume that they are bounded over finite time. Then, if the initial conditions satisfy
\[
\begin{align*}
u^2(0,x) - u^1(0,x) &\leq \varepsilon, \quad \text{for all } x, \quad (6.2.47) \\
u^2(t,x) - u^1(t,x) &\leq \varepsilon e^t, \quad \text{for all } t \text{ and } x. \quad (6.2.48)
\end{align*}
\]
The same holds for the absolute values of the differences.

Proof. Set \( v = u^2 - u^1 \). Then
\[
\partial_t v = \frac{1}{2} \partial_x^2 v + \frac{F(u^2) - F(u^1)}{u^2 - u^1} v .
\]
By assumption on \( f \), the coefficient of \( v \) is at most 1. Now let \( v^+ \) solve (6.2.49) with initial conditions
\[
\begin{align*}
v^+(0,x) &\leq v(0,x) \lor 0,
\end{align*}
\]
then by the maximum principle, for all times \( t \geq 0 \), \( v^+(t,x) \geq v(t,x) \lor 0 \) remains true. Next let \( \bar{v} \) be a solution of the linear equation
\[
\partial_t v = \frac{1}{2} \partial_x^2 v + v
\]
with initial condition \( \bar{v}(0,x) = \varepsilon \). Clearly, \( \bar{v}(t,x) = \varepsilon e^t \). On the other hand, by the maximum principle and the assumption on the initial conditions, \( v^+(t,x) \leq \bar{v}(t,x) \). Hence (6.2.48) follows. The same argument can be made for the negative part of \( v \). \( \square \)

An application of this lemma is the next corollary.

Corollary 6.14 ([22]). Let \( u \) be a solution of the F-KPP equation satisfying the standard conditions and assume that
\[
u(t,x + m(t)) \to w^λ(x), \quad \text{as } t \uparrow \infty,
\]
uniformly in \( x \). Then,
\[
m(t + s) - m(t) \to \lambda s,
\]
uniformly for \( s \) in compact intervals. Hence, \( m(t)/t \to \lambda \).

Proof. Since \( w^λ \) is stationary, the previous lemma implies that for all \( t \geq 0 \) and \( s_0 \geq 0 \), uniformly in \( x \leq s_0 \),
\[
\sup_{x \in \mathbb{R}} \left| u(t+s,x+\lambda s + m(t)) - w^λ(x) \right| \leq e^{s_0} \left| u(t,x + m(t)) - w^λ(x) \right|,
\]
where the right-hand side tends to zero by assumption. The next lemma states a quantitative comparison result. Thus
\[
\sup_x \left| u(t+s,x+\lambda s + m(t)) - u(t+s,x+m(t+s)) \right| \to 0,
\]
as \( t \uparrow \infty \), which implies \( m(t+s) - m(t) \to \lambda s \). \( \square \)
Lemma 6.15 (Lemma 3.3 in [22]). Let \( u^i, i = 1, 2, \) be solutions of the F-KPP equation with \( F'(u) \leq 1. \) If \( 0 \leq u^i(0, x) \leq 1, \) for all \( x, \) and \( x > 0 \)
\[
u^i(0, x) = u^2(0, x),
\]
then there exists a constant, \( C, \) such that
\[
|u^2(t, x) - u^1(t, x)| \leq Ct^{-1/4},
\]
for \( x > \sqrt{2t} - 2^{-5/2} \ln t.\)

Proof. The proof is similar to that of Lemma 6.12. With the same definitions as in that proof, we see that, by the Feynman-Kac formula, the solution \( \tilde{v} \) of the linear equation is given as
\[
\tilde{v}(t, x) = e^{t \sqrt{2/\pi}} \int_{-\infty}^{0} e^{-\frac{(x-y)^2}{2t}} dy \leq \frac{e^t}{\sqrt{2\pi x/\sqrt{t}}} e^{-\frac{x^2}{2t}}.
\]
For \( x = \sqrt{2t} - 2^{-5/2} \ln t, \) the right hand side is bound asymptotically equal to \( t^{-1/4}/\sqrt{2\pi}. \) \( \Box \)

The next lemma is the first result linking the behaviour of the tail of the initial distribution to the convergence to a travelling wave.

Lemma 6.16 (Lemma 3.4 in [22]). Let \( u \) be a solution of the F-KPP equation satisfying the standard setting. If \( \lambda = \sqrt{2} \) assume, in addition, that \( 1 - F'(u) = O(u^\rho) \) for some \( \rho > 0. \) If the initial condition is such that, for some \( \lambda \geq \sqrt{2}, \) and a function \( \gamma(x) \uparrow 1, \) as \( x \uparrow \infty, \)
\[
u(0, x) = \gamma(x) w^\lambda(x),
\]
where \( w^\lambda \) is a solution of the stationary equation with speed \( \lambda, \) then
\[
\sup_{x \geq b_\lambda(t)} \left| u(t, x + \lambda t) - w^\lambda(x) \right| \to 0, \quad as \ t \to \infty,
\]
where \( b_\lambda(t) = (\sqrt{2} - \lambda)t - \frac{1}{8} \ln t. \)

Proof. Set
\[
u^N(0, x) \equiv \begin{cases} u(0, x), & \text{for } x > N, \\ w^\lambda(x), & \text{for } x \leq N. \end{cases}
\]
We know that \( w^\lambda \) decays exponentially for \( x \to \infty. \) So due to (6.2.58), for any \( \delta >, \) there is \( N < \infty, \) such that for all \( x \in \mathbb{R}, \)
\[
w^\lambda(x + \delta) \leq u^N(0, x) \leq w^\lambda(x - \delta).
\]
The maximum principle implies then that
\[ w^\lambda(x + \delta) \leq u^N(t, \lambda t + x) \leq w^\lambda(x - \delta). \] (6.2.62)

On the other hand, for \( x \geq b(t) + N \), \( |u^N(t,x) - u(t,x)| \leq Ct^{-1/4} \). Thus at the expense of a little error of order \( t^{-1/4} \), we can replace \( u^N \) by \( u \) in (6.2.62) for such \( x \). But this gives (6.2.59), since \( w^\lambda \) is uniformly continuous. \( \square \)

The next proposition further strengthens the point that if a solutions approaches the travelling wave for large \( x \), then it does so also for smaller \( x \). This will be one main tool to turn the tail asymptotics of Theorem ?? into a proof of convergence to the travelling wave.

**Proposition 6.17 (Proposition 3.3 in [22]).** Under the assumptions of Lemma 6.16, if for some \( N \), for all \( x \geq N \),

\[ \gamma_1^{-1}(x)w^\lambda(x) - \gamma_2(t) \leq u(t,x + m(t)) \leq \gamma_1(x)w^\lambda(x) + \gamma_2(t), \] (6.2.63)

where \( \gamma_1(x) \to 1 \), as \( x \to \infty \), and \( \gamma_2(t) \to 0 \), as \( t \to \infty \). Then there exists \( c(t) \) tending to \(-\infty\) as \( t \to \infty \), such that

\[ \sup_{x \geq c(t)} |u(t,x + m(t)) - w^\lambda(x)| \to 0, \quad \text{as } t \to \infty \] (6.2.64)

**Proof.** The proof of this proposition is a relatively simple application of the maximum principle via the preceding lemma.

Some nice property hold if \( F \) not only satisfy the standard conditions, but it will even be concave, i.e. \( F(u)/u \) will be decreasing. This holds if \( F \) comes from BBM with binary branching, where \( F(u)/u = (1-u) \). In other situations one can compare solutions with those corresponding to upper and lower concave hulls,

\[ F^+(u) \equiv u \max(F(u')/u', u' \geq u), \] (6.2.65)

\[ F^-(u) \equiv u \max(F(u')/u', u' \leq u). \] (6.2.66)

Then the following lemma can be applied.

**Lemma 6.18 (Lemma 3.5 in [22]).** Assume that \( F \) is concave and \( u^i, i = 1, 2, 3 \) are solutions of the F-KPP equation satisfying the standard conditions. Then,

\[ u^3(0,x) \leq u^2(0,x) + u^1(0,x), \quad \text{for all } x, \] (6.2.67)

implies

\[ u^3(t,x) \leq u^2(t,x) + u^1(t,x), \quad \text{for all } x \text{ and for all } t \geq 0. \] (6.2.68)

Similarly, for any \( 0 < c \leq 1 \), if

\[ u^1(0,x) \geq cu^2(0,x), \] (6.2.69)

for all \( x \), then
6.2 The maximum principle and its applications

\[ u^1(t,x) \geq cu^2(t,x), \quad (6.2.70) \]

holds for all \( t \geq 0 \) for all \( x \).

The proof is a straightforward application of the maximum principle.

The following proposition compares the convergence of general initial data to the Heaviside case that was studied in Kolmogorov's theorem.

**Proposition 6.19 (Proposition 3.4 in [22])**. Let \( u \) be a solution of the F-KPP equation that satisfies the standard conditions with concave \( F \). Assume that for some \( t_0 \geq 1 \) and \( \eta > 0 \),

\[ u(t_0,0) \geq \eta. \quad (6.2.71) \]

Then for all \( 0 < \theta < 1 \) and \( \epsilon > 0 \), there exists \( T \) (independent of \( \eta \)) and a constant \( c_\theta \), such that for \( t \geq T \),

\[ u(t + t_0 - \theta^{-1} \ln \eta, x) > w_{\sqrt{2}}\left(|x| - m_{1/2}(t) + e^\theta\right) - \epsilon. \quad (6.2.72) \]

In particular, for all \( \delta > 0 \),

\[ u(t + t_0 - \theta^{-1} \ln \eta, x) \to 1, \quad (6.2.73) \]

uniformly in \( |x| \leq (\sqrt{2} - \delta)t \). If \( 1 - F(u) \leq u^0 \), then we may chose \( \theta = 1 \).

**Proof.** We need some preparation.

**Lemma 6.20 (Lemma 3.6 in [22])**. Assume that \( u(t_0,0) \geq \eta \), for some \( t_0 \geq 1 \) and \( \eta > 0 \). Then for either \( J = [0,1] \) or \( J = [-1,0] \),

\[ u(t_0,x) \geq \eta / 10, \quad \text{for all } x \in J. \quad (6.2.74) \]

**Proof.** We bound the solution by that of the linear F-KPP equation and use the Feynman-Kac representation. This implies that

\[ u(t_0,0) \leq e \int_\mathbb{R} u(t_0 - 1,y) \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy. \quad (6.2.75) \]

This is bounded from above by

\[ 2e \max \left( \int_{-\infty}^{0} u(t_0 - 1,y) \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy, \int_{0}^{\infty} u(t_0 - 1,y) \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy \right). \quad (6.2.76) \]

Let us say that the maximum is realised for the integral over the positive half line. Then use that, for \( y \in [0,1] \)

\[ \max_{y \geq 0} e^{(y-x)^2/2 - y^2/2} = e^{x^2/2} \leq e^{1/2}. \quad (6.2.77) \]

Thus
Now use that $\partial_t u \geq \partial^2_{xx} u$ to see that
\[
\int_{-\infty}^{\infty} u(t_0 - 1, y) e^{-(x-y)^2/2} \frac{dy}{\sqrt{2\pi}} \leq u(t_0, x).
\]
This gives (6.2.74). \qed

**Lemma 6.21.** Let $u(t, x)$ as usual and
\[
u(0, x) = \begin{cases} \eta, & x \in J, \\ 0, & x \notin J, \end{cases}
\]
for some interval $J$. For any $\theta \in (0, 1)$, there is $C^\theta_1 > 0$, such that for $t \leq \theta^{-1} \ln(C^\theta_1 / \eta)$, and all $x \in \mathbb{R}$,
\[
u(t, x) \geq \eta \theta^t \int_{J} e^{-(x-y)^2/(2t)} \frac{dy}{\sqrt{2\pi t}}.
\]
If (5.4.11) holds, there is $C_1 > 0$, such that for $t < -\ln \eta$,
\[
u(t, x) \geq C_1 \eta \epsilon^t \int_{J} e^{-(x-y)^2/(2t)} \frac{dy}{\sqrt{2\pi t}}.
\]

**Proof.** The key idea of the proof is to exploit the fact that if $u$ is small, then $F(u)/u$ is close to one. An upper bound then shows that for a short time, will grow as claimed. We introduce a comparison solution $u^\theta$ with the same initial conditions as $u$ and with $F(u)$ replaced by $F^\theta(u) \equiv F(u) \wedge (\theta u)$. Clearly $F^\theta(u) \leq F(u)$ and so, by the maximum principle, $\nu(t, x) \geq u^\theta(t, x)$, for all $t \geq 0$ and $x \in \mathbb{R}$. On the other hand,
\[
\partial_t u^\theta \leq \frac{1}{2} \partial^2_{xx} u^\theta + \theta u^\theta,
\]
we have the trivial upper bound
\[
u^\theta(t, x) \leq \eta \theta^t,
\]
for all $t \geq 0$ and $x \in \mathbb{R}$. Since $F'(u) \leq 1$, for given $\theta < 1$, there will be a $C^\theta_1 > 0$, such that for all $0 \leq u \leq C^\theta_1$, $F(u)/u \leq \theta$, and hence $F^\theta(u)/u = \theta$. By the bound (6.2.84), this is true at least up to time $t^* = \theta^{-1} \ln(C^\theta_1 / \eta)$. This gives the claimed lower bound. In the case when (5.4.11) holds, we can choose $\theta = 1$ and $u^\theta = u$. Then we know that $F(u(t, x))/u(t, x) \geq 1 - C^\theta_2 (\eta \epsilon^t)^\rho$, which we can use up to $t^* = -\ln \eta$. Using again the Feynman-Kac formula, we get for $t < t^*$,
\[
u(t, x) \geq e^{-\int_{t^*}^t (1-C^\theta_2 (\eta \epsilon^s)^\rho) ds} \int_{J} e^{-(x-y)^2/(2s)} \frac{dy}{\sqrt{2\pi s}}.
\]
We now prove the proposition. We consider only the case when (5.4.11) holds and thus set \( \theta = 1 \). Set \( t_1 = -\ln \eta, \ t_2 = t_0 + t_1 \). The two preceding lemmas give

\[
\int \frac{e^{-(x-y)^2/(2t)}}{\sqrt{2\pi t}} \, dy,
\]

which is what is claimed. \( \square \)

Now set \( x = z + m_{1/2}(t) \). Then the right hand side of (6.2.90) converges to

\[
C_4 \left( w^{\sqrt{2}}(z) - w^{\sqrt{2}}(z+1) \right).
\]

Using the tail behaviour of \( w^{\sqrt{2}} \) this is for large \( z, w^{\sqrt{2}}(z+c_1) \), for some \( 0 < c_1 < \infty \), and hence

\[
u(t,z+m_{1/2}(t) - c_1) \geq \gamma_1(z)w^{\sqrt{2}}(z) - \gamma_2(t).
\]

Then for all \( \varepsilon > 0 \) for large enough \( t \),

\[
u(t,z+m_{1/2}(t) - c_1) \geq w^{\sqrt{2}}(z) - \varepsilon,
\]

if \( z \geq c(t) \), where \( c(t) \) is from Proposition 6.17 and tends to \(-\infty\) as \( t \uparrow \infty \). This also implies that for large enough \( t \),

\[
u(t,c(t) + m_{1/2}(t) - c_1) \geq w^{\sqrt{2}}(c(t)) - \varepsilon/2 \geq 1 - \varepsilon.
\]

Since \( \nu(t,x) \) is decreasing in \( x \) for \( x \geq 0 \), we have that for \( 0 \leq x \leq m_{1/2}(t) + c(t) - c_1 \),

\[
u(t,x) \geq 1 - \varepsilon \geq w^{\sqrt{2}}(x - m_{1/2}(t) + c_1) - \varepsilon.
\]

On the other hand, for \( x \geq m_{1/2}(t) + c(t) - c_1 \), it follows from (6.2.93) that
\[ u^1(t, x) \geq w^{\sqrt{2}}(x - m_{1/2}(t) + c_1) - \varepsilon. \] (6.2.96)

This implies that for all \( x \geq 0 \),
\[ u(t + t_0 - \ln \eta + 1, x) \geq u^1(t, x) \geq w^{\sqrt{2}}(x - m_{1/2}(t) + c_1) - \varepsilon. \] (6.2.97)

This can be written as
\[ u(t + t_0 - \ln \eta, x) \geq u^1(t, x) \geq w^{\sqrt{2}}(x - m_{1/2}(t) + c) - \varepsilon, \] (6.2.98)

with a slightly changed constant \( c \). The case when \( x < 0 \) follows similarly.

Finally, since \( m_{1/2}(t) \sim \sqrt{2}t \) and \( w^{\sqrt{2}}(z) \uparrow 1 \), as \( z \downarrow -\infty \), it follows that
\[ u(t + t_0 - \ln \eta, x) \uparrow 1, \] (6.2.99)

uniformly in \( t \) if \( |x| \leq (\sqrt{2} - \delta)t \), which is (??). \( \square \)

The final proposition of this section asserts that the scaling function \( m_{1/2}(t) \) is essentially concave.

**Proposition 6.22 (Proposition 3.5 in [22])**. Let \( u \) be a solution of the F-KPP equation satisfying the standard conditions and Heaviside initial conditions. Assume further that \( F \) is concave. Then there exists a constant \( M > 0 \), such that for all \( s \leq t \),
\[ m_{1/2}(s) - \frac{s}{t} m_{1/2}(t) \leq M. \] (6.2.100)

**Proof.** Eq. (6.2.100) will follow from the fact that, for all \( t_1 \leq t_2 \) and \( s \geq 0 \),
\[ m_{1/2}(t_1 + s) - m_{1/2}(t_1) \leq m_{1/2}(t_2 + s) - m_{1/2}(t_2) + M. \] (6.2.101)

Namely, assume that (6.2.100) fails for some \( s \leq t \). Define
\[ T = \inf\{ r \in [s, t] : m_{1/2}(r)/r = \alpha \}, \] (6.2.102)

where \( \alpha = \inf_{q \in [s, t]} m_{1/2}(q) \). Let \( t_1 = \sup\{ r \in [0, s] : m_{1/2}(r)/r = \alpha \} \). Put \( t_2 = s \), and \( S = T - s \). Then
\[ m_{1/2}(t_2) > \alpha t_2 + M, \quad \text{and} \quad m_{1/2}(t_2 + S) = \alpha(t_2 + S). \] (6.2.103)

Hence
\[ m_{1/2}(t_2 + S) - m_{1/2}(t_2) < \alpha S - M. \] (6.2.104)

Furthermore, \( m_{1/2}(t_1) = \alpha t_1 \), and \( m_{1/2}(t_1 + S) \geq \alpha(t_1 + S) \), so that
\[ m_{1/2}(t_1 + S) - m_{1/2}(t_1) \geq \alpha S. \] (6.2.105)

(6.2.104) and (6.2.105) contradict (6.2.101).

It remains to prove (6.2.101). But by Corollary 6.2.32 for \( t_2 \geq t_1 \),
6.3 Estimates on solutions of the linear F-KPP equation

The linear F-KPP equation,
\[ \partial_t u = \frac{1}{2} \partial^2_x u + u, \quad (6.3.1) \]
has already served as reference in the previous sections. We derive now some precise estimates on the behaviour of its solutions from the explicit representation of its solution as
\[ \phi(t,x) = \frac{e^t}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} u(0,y) e^{-\frac{(x-y)^2}{2t}} dy. \quad (6.3.2) \]

This representation reduces the work to get good bounds into standard estimates on Gaussian integrals, i.e. the Gaussian tail bounds (1.2.11) and cutting up the range of integration.

Recall that \( \lambda, b \) are in the relation \( b = \lambda - \sqrt{\lambda^2 - 2} \).

We first prove an upper bound.

**Lemma 6.23 (Lemma 4.1 in [22]).** Assume that for some \( h > 0 \) and for some \( x_0 \), for all \( x \geq t_0 \),
\[ \int_{\mathbb{R}} u(1+b) u(0,y) dy \leq e^{-bx}. \quad (6.3.3) \]

Then, for each \( \delta \geq 0 \), and a constant \( C \),
\[ e^{\frac{t}{2}} \int_{\mathbb{R}} u(0,y) e^{-\frac{(x-y)^2}{2}} dx \leq e^{-b(x-\lambda t) - \delta^2 t/3} \quad (6.3.4) \]
for all \( t > 1 \) and for all \( x \), where
\[ \lambda = \frac{1}{b} + \frac{b}{2}. \]  

(6.3.5)

and \( J = (x - (b + \delta)t, x - (b - \delta)t) \).

**Proof.** The proof of this bound is elementary. See [22]. \( \square \)

Bramson states a variation on this bound as

**Corollary 6.24 (Corollary 1 in [22]).** Assume that for \( h > 0 \), for some \( 0 < b \leq \sqrt{2} \),

\[ \limsup_{x \to \infty} x^{-1} \ln \left( \int_{x}^{x+b} u(0, y) dy \right) \leq -b. \]  

(6.3.6)

Then, for \( \delta \geq 0 \) and \( 0 < \epsilon < b \),

\[ e^{t} \int_{t}^{\infty} u(0, y) \frac{e^{-\frac{(x-y)^2}{2t}}}{\sqrt{2\pi t}} dy \leq e^{-b(x-\lambda t) + \epsilon x - \frac{\delta^2 t}{4}}, \]  

(6.3.7)

for all \( x \) and \( t \) large enough. \( J \) is as in the previous lemma. In particular,

\[ \ln \phi(t, x) \leq -b(x-\lambda t) + \epsilon x. \]  

(6.3.8)

We need a sharper bound on \( \phi(t, x) \).

**Lemma 6.25.** Under the hypothesis of Lemma 6.23, for all \( x \),

\[ \phi(t, x) \leq \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{b_0-x} e^{-\frac{y^2}{2t}} dy + K \frac{1}{\sqrt{2\pi t}} \int_{0-x}^{\infty} e^{-by^2} dy. \]  

(6.3.9)

In particular, for \( x \geq \lambda t \),

\[ \phi(t, x) \leq \frac{e^{-\frac{x^2}{2}}}{{\sqrt{2\pi x}}} + Ke^{-(x-\lambda)^2}. \]  

(6.3.10)

**Proof.** We shift coordinates so that the integral in (6.3.2) takes the form \( \int_{-\infty}^{\infty} u(x+y) \exp(-y^2/(2t)) dy \). We split the integral in (6.3.2) into the part from \( \inf t \) to \( t_0 - x \) and the rest. In the first part, we bound \( u \) by 1. In the second, we split the integral into pieces of length, i.e. we write

\[ \int_{t_0-x}^{\infty} u(0,x+y)e^{-y^2/(2t)} dy = \sum_{n=0}^{\infty} \int_{t_0-x}^{n+1+b_0-x} u(0,x+y)e^{-y^2/(2t)} dy \]  

(6.3.11)

\[ \leq \sum_{n=0}^{\infty} e^{-\min_{\mid x+q_0-x+n+1+b_0-x\mid} (n+1+b_0-x)} u(0,x+y) dy \]

\[ \leq \sum_{n=0}^{\infty} e^{-b(n+1+b_0)} e^{-(n+1+b_0)^2/(2t)}. \]
6.3 Estimates on solutions of the linear F-KPP equation

We would like to reconstruct the integral from the last line. To do this, note that, for any \( n \),

\[
\max_{y \in [n+t_0-x, n+1+t_0-x]} y^2/(2t) - \min_{y \in [n+t_0-x, n+1+t_0-x]} y^2/(2t) \leq (2(n + t_0 - x \pm 1) + 1)/2t,
\]

(6.3.12)

where \( \pm \) depends on whether \( n + t_0 - x \) is positive or negative. For all \( n \) such that \( 2|n + t_0 - x| \leq Ct \), this is bounded by a constant, while otherwise, summands are anyway smaller than \( \exp(-Ct) \) which are negligible is \( C \) is chosen large enough.

Hence

\[
\sum_{n=0}^{\infty} e^{-\min_{y \in [n+t_0-x, n+1+t_0-x]} y^2/(2t)} e^{-b(n+t_0)} \leq e^{-bx} \int_{t_0-x}^{\infty} e^{-by^2/(2t)} dy.
\]

(6.3.13)

From here (6.3.9) follows. The last integral is bounded from above by

\[
e^{-bx+b^2t/2\sqrt{2\pi}}.
\]

(6.3.14)

Combining this with a standard Gaussian estimate for the first integral in (6.3.9) and recalling the definition of \( \lambda \) gives (6.3.10).

The next lemma provides a lower bound.

**Lemma 6.26.** Assume that for some \( h > 0 \),

\[
\liminf_{x \to \infty} x^{-1} \ln \left( \int_{x}^{x(1+h)} u(0, y) dy \right) \geq -b,
\]

(6.3.15)

with \( 0 < b < \sqrt{2} \). Let \( M > \lambda \) be fixed. Then, for all \( \varepsilon > 0 \), there is \( T > 0 \) such that, for \( (t, x) \in A_T \) and \( x \leq Mt \),

\[
\ln \phi(t, x) \geq b(x - \lambda t) - \varepsilon x,
\]

(6.3.16)

where \( A_T = \{(t, x) : t \geq T, x \geq b_0 t\} \cup \{(t, x) : 1 \leq t \leq T, x \geq b_0 t\} \).

**Corollary 6.27.** Let \( m^\phi(t) \) be defined by \( \phi(t, m^\phi(t)) = 1/2 \). If

\[
\lim_{x \to \infty} x^{-1} \ln \left( \int_{x}^{x(1+h)} u(0, y) dy \right) = -b,
\]

(6.3.17)

for \( 0 < b < \sqrt{2} \). Then

\[
\lim_{t \to \infty} m^\phi(t)/t = \lambda.
\]

(6.3.18)

If the initial data satisfy (6.3.6) with \( b = \sqrt{2} \), then \( \lambda = \sqrt{2} \).

**Proof.** In the first case, the result follows from the bounds on \( \phi \) in Lemma 6.26 and Corollary 6.24. In the second case, the upper bound is still valid with \( \lambda = \sqrt{2} \), while the lower bound follows by direct computation. \( \square \)
6.4 Brownian bridges

The use of the Feynman-Kac representation for analysing the asymptotics of solutions of the F-KPP equation requires a rather detailed control on the behaviour of Brownian bridges.

The following lemma, due to Bramson, gives some basic estimates.

Lemma 6.28 (Lemma 2.2. in [22]). Let \(\xi\) be a Brownian bridge from 0 to 0 in time \(t\). Then the following estimates hold.

(i) For \(y_1, y_2 > 0\),

\[
\mathbb{P}\left( \exists_{0 \leq s \leq t} : \xi(s) \geq (sy_1 + (t-s)y_2)/t \right) \leq e^{-2y_1y_2/t}. \tag{6.4.1}
\]

(ii) For \(x > 0\) and \(0 < s_0 < t\),

\[
\mathbb{P}\left( \exists_{0 \leq s \leq s_0} : \xi(s) \geq x \right) \leq \frac{2\sqrt{s_0}}{x} e^{-x^2/(2s_0)}. \tag{6.4.2}
\]

(iii) For \(y_1, y_2 > 0\), and \(0 \leq x \leq y_2/2\),

\[
p(t,x;\forall 0 \leq s \leq t : \xi(s) \leq y_1) \geq \sqrt{\frac{2}{\pi t}} y_1 e^{-x^2/(2t)} \left( x - (y_1 + 2y_2) e^{1/2} - y_1(x+y_1)^2 / t \right),
\]

where \(p(t, x; A)\) denotes the density of the Brownian motion restricted to the event \(A\).

(iv) For \(y_1, y_2 > 0\), and \(c_0\), let \(\xi_{y_1, y_2}\) be a Brownian bridge from \(y_1\) to \(y_2\) in time \(t\). Then

\[
\mathbb{P}\left( \forall 0 \leq s \leq t : \xi_{y_1, y_2}(s) > c_0(s \wedge (t-s)) \right) = \sqrt{\frac{2}{\pi t}} \int_0^\infty \left( 1 - e^{-4y_2t} \right) \left( 1 - e^{-4y_1t} \right) \times e^{-2(x+c_0/2-(y_1+y_2)/2)^2 / t} \, dx. \tag{6.4.4}
\]

**Proof.** The proof is a nice exercise using the reflection principle. We first prove (i). Fix \(t\) and let \(\xi\) be a Brownian bridge in time \(t\). Then the probability of BM to end in \(y_2 + y_1\) and to cross the level \(y_2\) is the same as the probability to stop in \(y_2\) and to cross the level \(y_1\), hence (6.4.5) equals

\[
p(t, y_2 + y_1) p(t, y_2 - y_1) = e^{-2y_1y_2/t}, \tag{6.4.6}
\]

where \(p(t, x)\) is the heat kernel. This proves (i). To prove (ii), Bramson uses that

\[
\mathbb{P}\left( \exists_{0 \leq s_0} : \xi(s) > x \right) \leq 2 \mathbb{P}\left( \exists_{0 \leq s_0} : B_s > x \right). \tag{6.4.7}
\]

(which I don’t quite see how to get easily). Since the law of the Brownian bridge at time \(t/2\) is the same as that of \(B_{t/2}\), the probability of the bridge to exceed a level \(x\).
6.4 Brownian bridges

Fig. 6.1 Use of the reflection principle

on an interval \([0, s_1]\) for \(s_1 \leq t/2\) is the same as that of Brownian motion. This gives the result if \(s_0 \leq t/2\). If \(s_0 > t/2\), we have

\[
P(∃s ≤ s_0 : z_t(s) > x) ≤ P(∃s ∈ [t/2, s_0] : z_t(s) > x) + P(∃s ∈ [0, t/2] : z_t(s) > x).
\]

(6.4.8)

For the second probability, we have that

\[
P(∃s ∈ [t/2, s_0] : z_t(s) > x) = P(∃s ∈ [t−s_0, t/2] : z_t(s) > x) \leq P(∃s ∈ [0, t/2] : z_t(s) > x).
\]

(6.4.9)

As both probabilities are the same as that for Brownian motion, we get a rather crude bound with the factor two in (6.4.7).

Applying the reflection principle to the Brownian motion probability and using the standard Gaussian tail asymptotics then yields the claimed estimate. To prove (iii), one uses the reflection principle to express the restricted density as

\[
p(t, x; ∀0 ≤ s ≤ t : z_t(s) \in [−y_1, y_2]) = p(t, x) − p(t, x + 2y_1) − p(t, x − 2y_1) + p(t, x + 2y_1 − 2y_2) + p(t, x − 2y_1 − 2y_2).
\]

(6.4.10)

Inserting the explicit expression of the heat kernel, this equals

\[
\frac{1}{\sqrt{2\pi t}} \left( e^{-x^2/2t} − e^{-(x+2y_1)^2/2t} − e^{-(x-2y_1)^2/2t} + e^{-(x-2y_2)^2/2t} \right)
\]

(6.4.11)

\[
= \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \left( 1 − e^{-2y_1(x+y_1)/t} − e^{-2y_1(y_1+y_2)} + e^{-2y_1(y_2-y_1)/t} \right)
\]

\[
= \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \left( 1 − e^{-2y_1(x+y_1)/t} − e^{-2y_1(y_2-y_1)/t} \left( 1 − e^{2y_1(x-2y_2-y_1)/t} \right) \right).
\]

Using that for \(u > 0\), \(e^{-u} ≥ 1 − u\) and \(e^{-u} ≤ 1 − u + u^2/2\), the right-hand side is bounded from below by
Using the assumption \( x \leq y_2/2 \) to simplify this expression yields the assertion of (iii). We skip the proof of part (iv) which is an application of the estimate (i). \[ \square \]

**Lemma 6.29.** Let \( \frac{1}{2} < \gamma < 1 \). Let \( z_t \) be a Brownian bridge from 0 to 0 in time \( t \). Then, for all \( \varepsilon > 0 \), there exists \( r \) large enough such that, for all \( t > 3r \),

\[
\mathbb{P} \left( \exists s \in [r, t-r] : |z_t'(s)| > (s \wedge (t-s))^{\gamma} \right) < \varepsilon. \tag{6.4.13}
\]

More precisely,

\[
\mathbb{P} \left( \exists s \in [r, t-r] : |z_t'(s)| > (s \wedge (t-s))^{\gamma} \right) < 8 \sum_{k=\lceil r \rceil}^{\lceil t/2 \rceil} k^{1/2} \gamma e^{-k^{2\gamma-1}/2}. \tag{6.4.14}
\]

**Proof.** The probability in (6.4.13) is bounded from above by

\[
\sum_{k=\lceil r \rceil}^{\lceil t/2 \rceil} \mathbb{P} \left( \exists s \in [k-1, k] : |z_t'(s)| > (s \wedge (t-s))^{\gamma} \right) \leq 2 \sum_{k=\lceil r \rceil}^{\lceil t/2 \rceil} \mathbb{P} \left( \exists s \in [k-1, k] : |z_t'(s)| > (s \wedge (t-s))^{\gamma} \right) \tag{6.4.15}
\]

by the reflection principle for the Brownian bridge. This is now bounded from above by

\[
2 \sum_{k=\lceil r \rceil}^{\lceil t/2 \rceil} \mathbb{P} \left( \exists s \in [0, k] : |z_t'(s)| > (k-1)^{\gamma} \right) \tag{6.4.16}
\]

Using the bound of Lemma 6.28 (ii) we have

\[
P \left( \exists s \in [0, k] : |z_t'(s)| > (k-1)^{\gamma} \right) \leq 4(k-1)^{\frac{1}{2} - \gamma} e^{-(k-1)^{2\gamma-1}/2}. \tag{6.4.17}
\]

Using this bound for each summand in (6.4.16) we obtain (6.4.14). Since the sum on the right-hand side of (6.4.17) is finite (6.4.13) follows. \[ \square \]
An important fact in the analysis of Brownian bridges is that hitting probabilities can be approximated by looking at discrete sets of times. This is basically a consequence of continuity of Brownian motion.

Let $\ell : [0, t] \to \mathbb{R}$ be a function that is bounded from above. Let $S_k \equiv \{s_{k,1}, \ldots, s_{k,2^k}\} \subset [0, t]$ be such that for all $k, j, s_{k,j} \in [(j-1)r2^{-k}, jr2^{-k}]$, and that
\[
\ell(s_{k,j}) \geq \sup\{\ell(s), s \in [(j-1)r2^{-k}, jr2^{-k}]\} - 2^{-k}.
\] (6.4.18)

Set $S^\alpha \equiv \bigcup_{k \leq n} S_k$, and set
\[
\ell^\alpha(x) = \begin{cases} 
\ell(x), & \text{if } s \in S^\alpha, \\
-\infty, & \text{else}.
\end{cases}
\] (6.4.19)

**Lemma 6.30** (Lemma 2.3 in [22]). With the notation above, the event $A \equiv \{z(s) > \ell(s), \forall 0 \leq s \leq t\}$ is measurable, and
\[
\lim_{n \to \infty} \mathbb{P}\left(z(s) > \ell^\alpha(s), \forall 0 \leq s \leq t\right) = \mathbb{P}(A).
\] (6.4.20)

We will not give the proof which is fairly straightforward.

### 6.5 Hitting probabilities of curves

We will need to compare hitting probabilities of different curves. In particular, we want to show that these are insensitive to slight deformations. The intuitive reason is related to what is called entropic repulsion: if a Brownian motion cannot touch a curve, it will actually stay quite far away from it, with large probability.

Let $L_t$ be a family of functions defined on $[0, t]$. For $C > 0$ and $0 < \delta < 1/2$, define
\[
L^\delta_t(s) = \begin{cases} 
L_t(s) - C_\delta s, & \text{if } 0 \leq s \leq t/2, \\
L_t(s) - C(t-s)\delta, & \text{if } s/2 \leq s \leq t,
\end{cases}
\] (6.5.1)
and
\[
L^\delta_t(s) = \begin{cases} 
L_t(s) + C_\delta s, & \text{if } 0 \leq s \leq t/2, \\
L_t(s) + C(t-s)\delta, & \text{if } s/2 \leq s \leq t.
\end{cases}
\] (6.5.2)

**Lemma 6.31** (Lemma 6.1 in [22]). Let $L_t$ be a family of functions such that for some $r_0 > 0$,
\[
L_t(s) \leq 0, \quad \forall s \in [r_0, t - r_0].
\] (6.5.3)
Then for all $\varepsilon > 0$ exists $r_\varepsilon$ large enough such that for all $r > r_\varepsilon$ and $t > 3r$
\[
\left| \frac{\mathbb{P}\left(L_t(s) > L^\alpha_t(s), \forall r \leq s \leq t-r\right)}{\mathbb{P}\left(L_t(s) > L^\alpha_t(s), \forall r \leq s \leq t-r\right)} - 1 \right| < \varepsilon.
\] (6.5.4)
Proof. In the proof of this lemma Bramson uses the Girsanov formula to transform the avoidance problem for one curve into that of another curve by just adding the appropriate drift to the BM. In fact is $γ$ is one curve and $\tilde{γ}$ is another, both starting and ending in 0, then clearly

$$P(\gamma'(s) > γ(s), \forall r \leq s \leq t - r) = P(\gamma'(s) - γ(s) + \tilde{γ}(s) > \tilde{γ}(s), \forall r \leq s \leq t - r).$$

(6.5.5)

But $\gamma'(s) - γ(s) + \tilde{γ}(s)$ is a Brownian bridge under the measure $\tilde{P}$, where $\tilde{P}$ is absolutely continuous w.r.t. to the law of BM with Radon Nikodym derivative

$$\frac{d\tilde{P}}{dP} = \exp\left(\int_0^t \partial_s \beta(s) d\gamma'_s - \frac{1}{2} \int_0^t (\partial_s \beta(s))^2 ds\right),$$

(6.5.6)

where $β(s) = \tilde{γ}(s) - γ(s)$.

We would be tempted to use this formula directly with $L_r$ and $L_t$, but the singularity of $s^{γ}$ would produce a diverging contribution. Since the condition on the Brownian bridge involves only the interval $(r, t - r)$, we can modify the drift on the intervals $[0, r]$ and $[t - r, t]$. Bramson uses

$$β_{r,t}(s) \equiv \begin{cases} 2C r^{δ-1} s, & \text{if } 0 \leq s \leq r, \\ 2C s^δ, & \text{if } r \leq s \leq t/2, \\ 2C(t - s)^δ, & \text{if } t/2 \leq s \leq t - r, \\ 2C r^{δ-1} (t - s), & \text{if } t - r \leq s \leq t. \end{cases}$$

(6.5.7)

Then the discussion above clearly shows that

$$P(\gamma'(s) > L_t(s), \forall r < s < t - r) = \mathbb{E}\left[e^{-\int_0^r \partial_s β_{r,t}(s) d\gamma'_s} \frac{1}{2} \int_0^r (\partial_s β_{r,t}(s))^2 ds \mathbb{1}_{\gamma'(s) > L_t(s), \forall r < s < t - r}\right].$$

(6.5.8)

Thus, to prove the lemma, we just need to show that the exponential tends to one. Now clearly

$$\int_0^t (\partial_s β_{r,t}(s))^2 ds = 8C^2 \left(2^{δ-1} + \frac{1}{2^{δ-1}} \left[\left(r/2\right)^{2δ-1} - r^{2δ-1}\right]\right),$$

(6.5.9)

which tends to zero as $r$ and $t$ tend to infinity. To treat the stochastic integral, we need an upper bound on the maximum of $\gamma'$. This is most easily done by sneaking in a condition $\{\gamma'(s) < 2C s^ε\}$, resp. $2C(t - s)^ε$, with $ε > 1/2$. It is easy to see that this condition is verified by the Brownian bridge with probability tending to one, so is makes no difference to add it. But then we get that

$$\int_0^t \partial_s β_{r,t}(s) d\gamma'_s = -\int_0^t \gamma'(s) \partial^2 \beta_{r,t}(s) ds.$$

(6.5.10)

Using now the bound on $\gamma'$, it is elementary to see that this integral is bounded in absolute value by $O(r^{δ+ε-1})$, which for, e.g. $ε = 1/2 + (1/2 - δ)/2$ tends to zero. \□
There are some further related results that relate hitting probabilities for certain other deformations of curves. If $L : [0,t] \to \mathbb{R}$, define, for $0 \leq \delta < 1/2$, $C > 0$,

$$
\theta_{x,t} \circ L(s) \equiv \begin{cases} 
L(s + x^\delta) + Cs^\delta, & \text{if } r \leq s \leq t/2, \\
L(s + (t-s)^\delta) + C(t-s)^\delta, & \text{if } t/2 \leq s \leq t-2r,
\end{cases} \quad (6.5.11)
$$

Then the following holds.

**Proposition 6.32 (Proposition 6.2 in [22])**. For all $\varepsilon > 0$ there exists $r_\varepsilon$ large enough such that for all $r > r_\varepsilon$ and all $t > 3r$,

$$
\frac{\left| \mathbb{P}(\bar{z}'(s) > (\theta_{x,t} \circ L_1(s)) \lor (\theta_{x,t}^{-1} \circ L_2(s)) \lor L_3(s), \forall r \leq s \leq t-r) - 1 \right|}{\mathbb{P}(\bar{z}'(s) > L_4(s), \forall r \leq s \leq t-r) - 1} < \varepsilon, \quad (6.5.12)
$$

and

$$
\frac{\left| \mathbb{P}(\bar{z}'(s) > L_4(s), \forall r \leq s \leq t-r) - 1 \right|}{\mathbb{P}(\bar{z}'(s) > \theta_{x,t}^{-1} \circ L_2(s), \forall r \leq s \leq t-r) - 1} < \varepsilon. \quad (6.5.13)
$$

Finally, Bramson proves the following monotonicity result:

**Proposition 6.33 (Proposition 6.3 in [22])**. Let $\ell_i, i = 1, 2$ be upper semicontinuous at all but finitely many values of $0 \leq s \leq t$, and let $\ell_1(s) \leq \ell_2(s)$, for all $0 \leq s \leq t$. Let $\bar{z}_{x,y}$ denote the Brownian bridge from $x$ to $y$ in time $t$. Then

$$
\mathbb{P}(\bar{z}_{x,y}(s) > \ell_2(s), \forall 0 \leq s \leq t) \geq \mathbb{P}(\bar{z}_{x,y}(s) > \ell_1(s), \forall 0 \leq s \leq t) \quad (6.5.14)
$$

is monotone increasing in $x$ and $y$.

### 6.6 Asymptotics of solutions of the F-KPP equation

We are now moving closer to proving Theorem 6.2. We still need some estimates on solutions. The first statement is a regularity estimate.

**Proposition 6.34 (Proposition 7.1 in [22])**. Let $u$ be a solution of the F-KPP equation satisfying the standard conditions. Assume that for some $h > 0$,

$$
\limsup_{r \to \infty} r^{-1} \ln \left( \int_t^{t(1+\varepsilon)} u(0,y)dy \right) \leq -\sqrt{2}. \quad (6.6.1)
$$

Then

$$
\lim_{r \to \infty} \sup_{0 \leq x_1 \leq x_2} (u(t,x_2) - u(t,x_1)) = 0. \quad (6.6.2)
$$

If, moreover, for some $\eta > 0, N, M > 0$, for $x < -M$, 
\[ \int_x^{x+N} u(0,y)dy \geq \eta, \quad (6.6.3) \]

then (6.6.2) holds also if the condition in the supremum is relaxed to \( x_1 \leq x_2 \).

**Proof.** Eq. (6.2.73) from Proposition 6.19 implies that for \( x \in [0,(\sqrt{2} - \delta)t] \), \( u(t,x) \to 1 \), for any \( \delta > 0 \). By the Feynman-Kac representation (6.1.2), using Condition (6.6.3), for all \( x < 0 \)

\[ u(t,x) = \begin{cases} u(0,x), & \text{if } x \leq \delta_1 s, \\ 0, & \text{if } x > \delta_1 s. \end{cases} \quad (6.6.5) \]

Now Lemma 6.12 implies that, for \( \delta_1 t \leq x_1 \leq x_2 \),

\[ u'(t,x_2) \leq u'(t,x_1). \quad (6.6.6) \]

If we set \( v'(t,x) = u(t,x) - u'(t,x) \), then

\[ v'(0,x) = \begin{cases} 0, & \text{if } x \leq \delta_1 s, \\ u(0,x), & \text{if } x > \delta_1 s. \end{cases} \quad (6.6.7) \]

But \( v' \) satisfies an equation

\[ \partial_t v' = \frac{1}{2} \partial_x^2 v' + k'(t,x)v', \quad (6.6.8) \]

where \( k' \) is some function of the solutions \( u, u' \). The important point is, however, that \( k'(t,x) \leq 1 \). Thus the Feynman-Kac representation yields the bound

\[ v'(t,x) \leq \frac{1}{\sqrt{2\pi t}} e^{-(x-x_1)^2/2t} \int_{\delta_1 s}^{\infty} u(0,y)dy. \quad (6.6.9) \]

Using the growth bound (6.6.1) on the initial condition, it follows from Corollary 6.24 that \( v'(t,x) \to 0 \), uniformly in \( x \geq t(\sqrt{2} - \delta^2/6) \). Using this together with (6.6.6) implies the desired convergence and proves the proposition. \( \square \)

The following propositions prepare for a more sophisticated use of the Feynman-Kac representation (6.1.6). This used in particular the comparison estimates of the hitting probabilities of different curves by the Brownian bridges. We use the notation
of the previous section and fix the reference curve
\[ L_{r_2}(s) = m_{1/2}(s) - \frac{s}{t}m_{1/2}(t)t^{-1}(t-s)\alpha(r,t), \] (6.6.10)
where \( m_{1/2} \) is defined in Eq. (6.2.28). \( \alpha(r,t) = \rho(t) \) will be specified later. Next we introduce the functions
\[ \mathcal{M}_r(s) \equiv \mathcal{M}_r(s) + \frac{s}{t}m_{1/2}(t) + \frac{s}{t} \alpha(r,t) = m_{1/2}(s) + \mathcal{M}_r(s) - L_{r_2}(s), \] (6.6.11)
and
\[ \mathcal{M}_r(s) \equiv \mathcal{M}_r(s) + \frac{s}{t}m_{1/2}(t) + \frac{s}{t} \alpha(r,t) = m_{1/2}(s) + \mathcal{M}_r(s) - L_{r_2}(s). \] (6.6.12)

At this point we can extend the rough estimate on \( m_{1/2}(t) \) given in (6.2.31) for Heaviside initial conditions to initial conditions satisfying (6.6.1).

**Lemma 6.35.** Assume that \( u(0,x) \) satisfies (6.6.1). Then (6.2.31) holds, i.e. \( m_{1/2}(t)/t \to \sqrt{2} \).

**Proof.** Recall that by the maximum principle, \( u(t,x) \leq \phi(t,x) \) with \( \phi \) defined as (6.3.2). On the other hand, it is fairly straightforward to show that, if the initial data satisfy (6.6.1), then \( m^{\phi}(t) \), i.e. the analog of \( m_{1/2}(t) \) defined for \( \phi \) behaves like \( \sqrt{2t} \). Hence \( \limsup_{t \to \infty} m_{1/2}(t)/t \leq \sqrt{2} \). The bound (6.2.73) in Proposition 6.19 implies, on the other hand, that \( \liminf_m m_{1/2}(t)/t \geq \sqrt{2} \). \( \square \)

The strategy of using the Feynman-Kac formula relies on controlling the term \( k(y,s) \) appearing in the exponent in appropriate regions of the plane, and then establishing probabilities that the Brownian bridge passes through those. The next proposition states that \( k \) is close to one above the curves \( \mathcal{M}_r \).

**Proposition 6.36 (Proposition 7.2 in [22]).** Let \( u \) be a solution to the F-KPP equation satisfying the standard assumptions, with initial data satisfying (6.6.1) for some \( h > 0 \). Then, there is a constant \( C_2 \) such that, for \( r \) large enough and \( y > \mathcal{M}_{r_2}(s) \), one has that
\[ k(s,y) \geq \begin{cases} 1 - C_2 e^{-\rho(s)\delta}, & \text{if } r \leq s \leq t/2, \\ 1 - C_2 e^{-\rho(t-s)\delta}, & \text{if } t/2 \leq s \leq t - r. \end{cases} \] (6.6.13)
Thus
\[ e^{3r-\delta} e^{\int_{s}^{t-r} k(t-s,x(s))ds} \to 1, \quad \text{as } r \to \infty, \] (6.6.14)
if \( x(s) > \mathcal{M}_{r_2}(t-s) \) for \( s \in [2r,t-r] \).

**Proof.** We give a proof of of this proposition under the slightly stronger hypothesis of Lemma 6.3.1. The strategy is the following:

(i) We prove (6.6.14) for \( \mathcal{M}_{r_2} \) defined with \( m_{1/2}(s) \) replaced by any function \( m(s) \) such that \( \sqrt{2s} \geq m(s) \geq \sqrt{2s} - C \ln s \), for any \( C < \infty \).
(ii) Use this bound to control \( u(t,x) \) for \( x \geq \sqrt{2t} - C \ln t \).
(iii) Deduce from this that
\[ m_{1/2}(t) \geq \sqrt{2t} - \frac{3}{2\sqrt{2}} \ln t + M, \]  
(6.6.15)
for \( M < \infty \).

We begin with Step (i) and assume that the \( m_{1/2} \) in the definition of \( \overline{M} \) is stated in item (i). By definition, for \( s \in [r, t/2] \), after some tedious computations,
\[ \overline{M}_{rs}(s) \geq m_{1/2}(s + s^\delta) + (4 + (\alpha(r, t) - m_{1/2}(t))/t)s^\delta. \]  
(6.6.16)

By Lemma 6.35 for large enough \( r < t \), the quantity in the bracket is positive and hence
\[ \overline{M}_{rs}(s) \geq m_{1/2}(s + s^\delta). \]  
(6.6.17)

We use that \( u(s, x) \leq \phi(s, x) \), and that by Lemma 6.25 under the Assumption (6.3.6) with \( b = \sqrt{2} \),
\[ \phi(s, m_{1/2}(s) + s^\delta + z) \leq e^{-(m_{1/2}(s + s^\delta))^2/(2s)} + e^{s\sqrt{2} - (m_{1/2}(s + s^\delta) - \sqrt{2}z)} \]
\[ \leq e^{s\sqrt{2} - \sqrt{2}\delta} + Ke^{-\sqrt{2}\delta}. \]  
(6.6.18)

Hence for \( y > \overline{M}_{rs}(s) \) and \( s \in [r, t/2] \),
\[ u(y, s) \leq C_3 e^{-\sqrt{2}\delta}. \]  
(6.6.19)

The analogous result follows in the same way for \( s \in [t/2, t - r] \). Using that \( F \) satisfies the standard conditions implies then (6.6.13). The conclusion (6.6.14) is obvious.

We now come to Step (ii).

**Lemma 6.37.** Let \( u \) be as in the proposition. Let \( x = x(t) = \sqrt{2t} - \Delta(t) \), where \( \Delta(t) \leq \sqrt{2t} - C \ln t \), for some \( C < \infty \). Then
\[ u(x, t) \geq C^{-1}e^{2\sqrt{2}\Delta(t)} \int_0^\infty u(0, y) \frac{e^{\sqrt{2}y - y^2/2t}}{\sqrt{2\pi t}} dy. \]  
(6.6.20)

**Proof.** We know that Step (i) that for our choices of \( \overline{M} < \overline{\overline{M}}_{s,s}(s) \) \( \overline{M}_{rs}(s) \) on \([2r, t - r] \),
\[ \exp \left( \int_{2r}^{t-r} k(s, \overline{\overline{M}}_{s,s}(s))ds \right) \geq C_3 e^r, \]  
(6.6.21)

where \( C_3 \) depends on \( r \) but not on \( t \). Now choose \( y_0 \) small enough such that \( \int_0^{y_0} u(0, y) dy > 0 \). Now we use the Feynman-Kac representation in the form (6.1.7) (with \( r = 0 \)) to get the bound
6.6 Asymptotics of solutions of the F-KPP equation

\[
    u(x,t) \geq \int_{\mathbb{R}} u(0,y) \frac{e^{-(x-y)^2/2t}}{\sqrt{2\pi t}} \mathbb{E} \left[ 1_{\{3_{x,y}(s) > \mathcal{M}_{st}(t-s), \forall s \in [2r,t-r] \}} e^{\int_{t-s}^{t} k(s,3_{x,y}(s)) \, ds} \right] \, dy
\]

\[
    \geq C_3 e^{t} \int_{y_0}^{\infty} u(0,y) \frac{e^{-(x-y)^2/2t}}{\sqrt{2\pi t}} \mathbb{P} \left( 3_{x,y}(s) > \mathcal{M}_{st}(t-s), \forall s \in [2r,t-r] \right) \, dy
\]

The curves $\mathcal{M}, \mathcal{M}$ can be chosen with $\alpha(r,t) = y_0$, with $y_0$ fixed (usually negative). One can always choose $m(s)$ such that $m(s) - \frac{1}{t} m(t) \leq M < \infty$. Then for large enough $r$, the probability in the last line can be bounded from below using our comparison results for hitting probabilities of Brownian bridges that imply that

\[
    \mathbb{P} (3_{x,y}(s) > \mathcal{M}_{st}(t-s), \forall s \in [r,t-r]) \geq C_1 \mathbb{P} (3_{x,y}(s) > \mathcal{M}_{st}(t-s), \forall s \in [r,t-r]) \geq C_1 \mathbb{P} (\exists s \in [r,t-r]),
\]

where in the last line $\exists \equiv \exists_0,0$. Using (5.2.7) and the fact that there is a positive probability that at times $r$ and $t - r$, the Brownian bridge is larger than, say, 1, one sees that there is a constant, $C$, depending only on $r$, such that

\[
    \mathbb{P} (\exists s \in [r,t-r]) \geq C / t.
\]

Hence we have that

\[
    u(t,x) \geq C_4 t^{-1} e^{t} \int_{y_0}^{\infty} u(0,y) \frac{e^{-(x-y)^2/2t}}{\sqrt{2\pi t}} \, dy
\]

Inserting $x = \sqrt{2t} - \Delta(t)$ and dropping terms that tend to zero, we arrive at (6.6.20). □

It is now easy to get the desired lower bound on $m_{1/2}$.

**Proposition 6.38 (Proposition 8.1 in [22]).** Under the hypothesis stated above, there exists a constant $C_0 < \infty$ and $t_0 < \infty$, such that for all $t \geq t_0$,

\[
    m_{1/2}(t) \geq \sqrt{2t} - \frac{3}{2\sqrt{2}} \ln t + C_0 \equiv n(s).
\]

**Proof.** We insert into (6.6.20) $\Delta(t) = \frac{D}{2\sqrt{2}} \ln t - z$. Then we see that

\[
    u(t,\sqrt{2t} - \Delta(t)) \geq C t^{-3/2} e^{\ln t} \int_{y_0}^{\infty} u(0,y) \frac{e^{\sqrt{2}y} e^{-y^2/2t}}{\sqrt{2\pi}} \, dy
\]

\[
    \geq C t^{-3/2} e^{D \ln t - z} \int_{y_0}^{A} u(0,y) e^{\sqrt{2}y} e^{-y^2/2t} \, dy
\]

\[
    \geq \tilde{C}(A, y_0) t^{-3/2} e^{-2\sqrt{2} z}.
\]

for some $A < \infty$ chosen such that $\int_{y_0}^{A} u(0,y) \, dy > 0$, where we just replaced the exponentials in the integral by their minimum on $[y_0, A]$. Clearly, this can be made
equal to 1/2 by setting $D = 3$ and choosing finite $z$ appropriately. This implies the assertion for $m_{1/2}$. □

Having shown that the true $m_{1/2}$ satisfies the hypothesis made in Step (i), we see that the assertion (6.6.13) holds for $\mathcal{M}$ defined with the true $m_{1/2}$. This concludes the proof of Proposition 6.36. □

**Remark 6.39.** Bramson proves this proposition under the weaker hypothesis (6.6.1). His argument goes as follows. First, by definition, $u(s + \sigma \delta, m_{1/2}(s + \delta)) = 1/2$. Next set $v(t, 0) \equiv u(s, m_{1/2}(s + \delta))$. Fix $s \ll 0$. Then $v(t, x)$ solves the F-KPP equation and $v(s, 0) = u(s, m_{1/2}(s + \delta))$. But by (6.6.21) in Proposition 6.19, if $v(t, 0) = \eta$, then $v(t + t_0 - \ln \eta, x) \uparrow 1$, as $t \to \infty$. But if $s^\delta + \ln \eta \uparrow \infty$, then $v(s^\delta, 0) \uparrow 1$, as $s^\delta$, in contradiction with the fact that $v(s^\delta, 0) = 1/2$. Hence we must have that $\ln \eta + s^\delta < \infty$, or $\eta \leq e^{-\sigma \delta}$.

This result already suggests that the contributions in the Feynman-Kac representations should come from Brownian bridges that stay above the curves $\mathcal{M}$ and thus enjoy the full effect of the $k$. To show this, one first shows that bridges staying below curves $\mathcal{M}$ do not contribute much.

Define the sets

$$G_{x,y}(r, t) \equiv \{3_{x,y} : \exists s \in [2r, t - r] : 3_{x,y}(s) \leq \mathcal{M}_{x,y}(t - s)\},$$

(6.6.28)

and

$$G_{x,y}(r_1, t) \equiv \{3_{x,y} : \exists s \in [2r \vee r_1, t - r_1] : 3_{x,y}(s) \leq \mathcal{M}_{x,y}(t - s)\}. \quad (6.6.29)$$

**Proposition 6.40** (Proposition 7.3 in [22]). Let $u$ be a solution that satisfies (6.6.1). Assume that $\alpha(r, t) = o(t)$ and that

$$m_{1/2}(s) \leq \begin{cases} \frac{1}{4} m_{1/2}(t) + \frac{1}{2} \alpha(r, t) + 8 s^\delta, & \text{if } s \in [r, t/2], \\ \frac{1}{4} m_{1/2}(t) + \frac{1}{2} \alpha(r, t) + 8(t - s)^\delta, & \text{if } s \in [t/2, t - r]. \end{cases} \quad (6.6.30)$$

Then for $y > \alpha(r, t)$ and $x \geq m_{1/2}(t)$, for $r$ large enough,

$$e^{3r - t} \mathbb{E}\left[\exp \left(\int_0^{t-r} \mathbb{1}_{[\mathcal{M}_{x,y}(s) \uparrow \mathcal{M}_{x,y}(t)]} ds\right)\right] \leq r^{-2} \mathbb{P}\left(G_{x,y}(r, t)\right). \quad (6.6.31)$$

**Proof.** The proof of this proposition is quite involved and will only be sketched. We need some localising events for the Brownian bridges. Set $I_j \equiv [j, j + 1) \cup (t - j - 1, t - j)$, for $j = 0, \ldots, j_0 - 1$, with $j_0 < t \leq j_0 + 1$, and $I_{j_0} = [j_0, t - j_0]$. Define the exit times

$$S_1(r, t) \equiv \sup \{s : 2r \leq s \leq t/2, 3_{x,y}(s) \leq \mathcal{M}_{x,y}(t - s)\} \quad (6.6.32)$$

$$S_2(r, t) \equiv \inf \{s : t/2 \leq s \leq t - r, 3_{x,y}(s) \leq \mathcal{M}_{x,y}(t - s)\} \quad (6.6.33)$$

and
Then define
\[ A_j(r,t) \equiv \{ 3_{x,y} : S(r,t) \not\subset I_j \}, \quad j = 0, \ldots, f_0. \] (6.6.35)
and further
\[ A_j^1(r,t) \equiv \{ 3_{x,y} \in A_{r,t} : 3_{x,y}(s) > -(s \wedge (t-s) + ys/t + x(t-s)/t, \forall s \in I_j \}. \] (6.6.36)

Finally, \( A^2_j(r,t) = A_j(r,t) - A^1_j(r,t) \).

First one shows that the probability not to cross below \( A \) in the interval \([r_1, t - r_1]\) is controlled by that not crossing in \([2r, t - r]\),
\[ \mathbb{P} \left( G^c_{x,y}(r_1; r, t) \right) \leq C_3 \alpha \frac{r_1}{r} \mathbb{P} \left( G^c_{x,y}(r, t) \right), \] (6.6.37)
for \( y \geq \alpha(r, t) \) and \( x \geq m_{1/2}(t) \). The proof of this [Lemma 7.1 in [22]] is via usual arguments for bridges (reflection principle).

The next step is more complicated. It consists in showing that if a bridge realises the event \( A^1_j(r,t) \) for large \( j \), i.e. if \( \mathcal{H}_{x,y} \) is crossed near the centre, then this will give a small contribution to \( u \). More precisely,
\[ e^{3r-t} \mathbb{E} \left[ e^{\int_0^{r_1} k(t-s, 3_{x,y}(s)) \, ds} A^1_j(r, t) \right] \leq C_4 e^{-j^8/4} \mathbb{P} \left( A^1_j(r, t) \right). \] (6.6.38)

The point is to that around the place where the curve \( \mathcal{H} \) is crossed, \( u(s, 3) \) is roughly \( u(s, m_{1/2} - s^\delta) \). Again by Proposition 6.19 at this position \( u \) will be very close to 1 at this point, and hence \( k \) will be very small. Fiddling around this produces a loss by a factor \( e^{-j^8/4} \).

Finally, the event in the left-hand side in (6.6.31) is decomposed into the union of the events \( A_j(r,t) \), and using the estimates (6.6.38), one obtains a bound of the form (6.6.31), where the choice of the factor \( 1/r^2 \) is quite arbitrary; in fact one could get any negative power of \( r \).

The next proposition is a reformulation of Proposition 6.32

**Proposition 6.41 (Proposition 7.4 in [22]).** Let \( u \) be as above and assume that (6.6.30) holds as in the previous proposition. For a constant \( M_1 > 0 \), assume that \( x \geq m_{1/2}(t) + M_1 \) and \( y \geq \alpha(r, t) + M_1 \). Then,
\[ \mathbb{P} \left( 3_{x,y}(s) > \mathcal{H}_{x,y}(t-s), \forall s \in [r, t-r] \right) \rightarrow 1, \] (6.6.39)
and
\[ \mathbb{P} \left( 3_{x,y}(s) > m_{1/2}(t-s), \forall s \in [r, t-r] \right) \rightarrow 1, \] (6.6.40)
uniformly in \( t \), as \( r \uparrow \infty \).
So far the probabilities concern only the bulk part of the bridges. To control the initial and final phases we need to slightly deform the curves. Define

$$M_{x,r,t}(s) = \begin{cases} M_{r,t}(s), & \text{if } s \in [0,t-2r], \\ (x + m_{1/2}(t))/2, & \text{if } s \in [t-2r,t], \end{cases} \quad (6.6.41)$$

$$M_{y,r,t}(s) = \begin{cases} M_{r,t}(s), & \text{if } s \in [r,t-2r], \\ y/2, & \text{if } s \in [0,r], \\ (x + m_{1/2}(t))/2, & \text{if } s \in [t-2r,t], \end{cases} \quad (6.6.42)$$

and

$$M'_{x,r,t}(s) = \begin{cases} M_{r,t}(s), & \text{if } s \in [r,t-2r], \\ -\infty, & \text{else}. \end{cases} \quad (6.6.43)$$

**Proposition 6.42 (Proposition 7.5 in [22])**. Let $u$ be as above and assume that (6.6.30) holds as in the previous proposition. Then,

$$\frac{\mathbb{P}(\exists_{s\in[0,t-2r]} M_{x,r,t}(s) > M_{x,y,r,t}(t-s), \forall s \in [0,t-r])}{\mathbb{P}(\exists_{s\in[0,t-r]} M'_{x,r,t}(s) > M'_{x,y,r,t}(t-s), \forall s \in [0,t-r])} \to 1, \quad (6.6.44)$$

uniformly in $t \geq 8r$, $x \geq x_0$, and $y \geq \alpha(r,t)$, as $r \uparrow \infty$. Moreover,

$$\frac{\mathbb{P}(\exists_{s\in[0,t]} M_{x,r,t}(s) > M_{x,y,r,t}(t-s), \forall s \in [0,t])}{\mathbb{P}(\exists_{s\in[0,t]} M'_{x,r,t}(s) > M'_{x,y,r,t}(t-s), \forall s \in [0,t])} \to 1, \quad (6.6.45)$$

and uniformly in $t \geq 8r$, $x \geq x_0$, and $y \geq \alpha(r,t)$, as $r \uparrow \infty$. Moreover, we will assume mostly that there exist $y_0$ such that

$$\int_{y_0}^{\infty} u(0,y)\,dy > 0. \quad (6.7.1)$$

Now set

$$\bar{y} = y - \alpha(r,t), \quad z = x - m_{1/2}(t), \quad \bar{z} = x - n(t). \quad (6.7.2)$$
We all the curves introduced above with an additional superscript \( n \) will denote these curves when \( m_{1/2}(s) \) is replaced by \( n(s) \) in their construction. We want an upper bound on the probability that a Brownian bridge lies above \( \mathbb{R} \).

**Lemma 6.43 (Corollary 1 in Chapter 8 of [22]).** Assume that \( \alpha(r,t) = O(\ln r) \). Then, if \( x \geq n(t) + 1, y \geq \alpha(r,t) + 1 \), there exists a constant \( C > 0 \), such that for \( r \) large enough,

\[
\mathbb{P}(\exists s, y(s) > \mathbb{R}(t-s), \forall s \in [2r, t-r]) \leq Cr \left( 1 - e^{-2\sqrt{y}/r} \right). 
\]

**(6.7.3)**

**Proof.** Playing around with the different curves, Bramson arrives at the upper bound

\[
C \mathbb{P}(\exists s, y(s) > 0, \forall s \in [2r, t-2r]).
\]

**(6.7.4)**

Now it is plausible that with positive probability, \( \exists \mathbb{E}^s \sim \mathbb{E}^s(t-2r) \sim \sqrt{r} \). Then the probability to stay positive is \( (1 - e^{-r/(t-4r)}) \), which for large \( t \) is about what is claimed.

We now move to the lower bound on \( m_{1/2} \).

**Proposition 6.44 (Proposition 8.2 in [22]).** Under the general assumptions of this section, for all \( x \geq m_{1/2}(t) + 1 \) and large enough \( t \),

\[
u(t, x) \leq C_2 e^{t} \int_{y_0}^{y} u(0, y) \frac{e^{-(x-y)^2/2t}}{2\sqrt{2\pi t}} \left( 1 - e^{-2\sqrt{y}/t} \right) dy, 
\]

**(6.7.5)**

where \( C_2 \) depends on initial data but not on \( t \). Hence

\[
u(t, x) \leq C_3 e^{-\sqrt{2t}},
\]

**(6.7.6)**

and

\[
m_{1/2}(t) \leq \sqrt{2t} - \frac{3}{2\sqrt{2}} \ln t + C_4,
\]

**(6.7.7)**

for some constant \( C_4 \).

**Proof.** Again we assume \( F \) concave. We may set \( y_0 = 0 \), and by the maximum principle, we may assume \( u(0, y) = 1 \), for \( y \leq 0 \). The main step in the proof of (6.7.5) is to use the Feynman-Kac representation and to introduce a one in the form \( 1_{G_{s,y}(r,t)} + 1_{\bar{G}_{s,y}(r,t)} \). Then we use Proposition 6.40 and Lemma 6.43. This yields that

\[
\mathbb{E} \left[ e^{-\int_0^t k(t-s, \mathbb{E}(s))} \right] \leq e^t \mathbb{P} \left( G_{s,y}(r,t) \right) + e^t r^{-2} \mathbb{P} \left( G_{s,y}(r,t) \right) \leq C_5 e^t \left( 1 - e^{-2\sqrt{y}/t} \right),
\]

**(6.7.8)**

for some constant \( C_5 \), for \( y \geq 1, x \geq n(t) + 1 \).

Using monotonicity arguments, one derives with some fiddling at the upper bound

\[
u(t, x) \leq C_6 e^{t} \int_{y_0}^{y} u(0, y) \frac{e^{-(x-y)^2/2t}}{2\sqrt{2\pi t}} \left( 1 - e^{-2\sqrt{y}/t} \right) dy,
\]

**(6.7.9)**
for \( x \geq m_{1/2}(t) + 1 \). To conclude we must show that \( \bar{z} \) can be replaced by \( z \) for which we want to show that \( n(t) \) and \( m_{1/2}(t) \) differ by at most a constant. Set \( z = x - \sqrt{2} t \). Then the left hand side of (6.7.9) equals

\[
C_6 e^{-\sqrt{2} z} \int_0^\infty u(0, y) e^{\sqrt{2} y} \frac{e^{-1/2(y - z)^2 / 2t}}{\sqrt{2 \pi t}} \left( 1 - e^{-2\bar{z}^2 / t} \right) dy \tag{6.7.10}
\]

\[
\leq C_6 e^{-1/2} \sqrt{\frac{2}{\pi}} \int_0^\infty y e^{\sqrt{2} y} u(0, y) dy \leq C_7 t^{-3/2} \bar{z} e^{-\sqrt{2} z} \leq C_8 \bar{z} e^{-\sqrt{2} z}, \tag{6.7.11}
\]

Since \( n(t) \leq m_{1/2}(t) \), this implies (6.7.6). But then

\[
u(t, m_{1/2}(t)) \leq C_8 (m_{1/2}(t) - n(t)) e^{-\sqrt{2}(m_{1/2}(t) - n(t))}. \tag{6.7.12}
\]

Since the right-hand side is bounded from below uniformly in \( t \), this implies that \( m_{1/2} - n(t) \) is bounded uniformly in \( t \), hence (6.7.7). \( \square \)

Proposition 6.44 gives the following upper bound on the distribution of the maximum of BBM that will be used in the next chapter.

**Lemma 6.45 (Corollary 10 in [4]).** There exists a numerical constants, \( \rho, t_0 < 0 \), such that for all \( x > 1 \), and all \( t \geq t_0 \),

\[
\mathbb{P} \left( \max_{0 \leq s \leq t} x_s(t) - m_{1/2}(t) \geq x \right) \leq \rho x \exp \left( -\sqrt{2} x - \frac{x^2}{2t} + \frac{3x}{2\sqrt{2} t} \right). \tag{6.7.13}
\]

**Proof.** For Heaviside initial conditions and using \( y_0 = -1 \), (6.7.5) gives

\[
u(t, x + m_{1/2}(t)) \leq C_2 e^{\frac{1}{t} \int_{-1}^0 \frac{e^{-(x + m(t) - y)^2 / 2t}}{\sqrt{2 \pi t}} \left( 1 - e^{2\bar{z}^2 / t} \right) dy}. \tag{6.7.14}
\]

Using that \( 1 - e^{-u} \leq u \), for \( u \geq 0 \) and writing out what \( m_{1/2}(t) \) is, the result follows. \( \square \)

Note that this bound establishes the tail behaviour for the probability of the maximum, Eq. (5.6.9), that was used in Section 5.6.

We are now closing in on the final step in the proof. Before we do that, we need one more a priori bound, which is the converse to (6.7.5).

**Corollary 6.46 (Corollary 1 in [22]).** Under the assumptions as before, for \( x \geq m_{1/2}(t) \) and all \( t \),

\[
u(t, x) \geq C_5 e^{\frac{1}{\alpha(r, t)} \int_0^\infty u(0, y) e^{-(x - y)^2 / 2t} \left( 1 - e^{2\bar{z}^2 / t} \right) dy}, \tag{6.7.15}
\]

if \( \alpha(r, t) = O(\ln r) \). \( C_5 \) does not depend on \( x \) or \( t \).
6.7 Convergence results

**Proof.** The main point is that we can bound the probability that the Brownian bridge stays above \( \mathcal{M}_{r,t} \) on \([2r, t−r] \) (up to constants) by the probability that \( \xi_{r} \) stays positive on \([r, t−r] \), which in turn is bounded by a constant times \((1−\exp(−2\sqrt{r}/t))\).

The rest of the proof is then obvious. \( \Box \)

By now we control our functions up to multiplicative constants. The reason for the errors is essentially that we do not control the integrals in the exponent in the Feynman-Kac formula well outside of the intervals \([2r, t−r] \). In the next step we remedy this. Now the representation (6.1.7) comes into play. This avoids having to control what happens up to time \( r \).

We now come to the final step, the proof of Theorem 6.2.

**Proof (of Theorem 6.2).** As advertised before, all reposes on the use of the Feynman-Kac representation in the form (6.1.7). Inserting a condition for the bridge to stay above a curve \( \mathcal{M}_{r,t} \), we get the lower bound

\[
\begin{align*}
\quad & u(t, x) \geq \int_{-\infty}^{\infty} u(r, y) \frac{e^{-(y-x)^2/(2(t-r))}}{\sqrt{2\pi(t-r)}} \\
\quad & \quad \times \mathbb{E} \left[ \exp \left( \int_{0}^{t-r} k(t-s, \xi_{r}(s)) ds \right) 1_{\xi_{r}(s) > \mathcal{M}_{r,t}(s), y \in [0, t-r]} \right] dy.
\end{align*}
\]

We want to show that condition on the path, the nonlinear term in the integral in the exponent can be replaced by 1. This is provided by Proposition 6.42 for the part of the integral between \( 2r \) and \( t−r \), since on this segment, \( \mathcal{M}_{r,t}(s) = \mathcal{M}_{r,t}(t−s) \). Now we are interested in \( x > m_{1/2}(t) \). For \( 0 \leq s \leq 2r \), \( \mathcal{M}_{r,t}(t−s) = (x + m_{1/2}(t))/2 \). Thus for \( x = m_{1/2}(t) + z \) with \( z \) large enough, by Proposition 6.44, for \( x(s) > \mathcal{M}_{r,t}(t−s) \),

\[
u(t, x(s)) \leq C_{1}z e^{-\sqrt{2}z}.
\]

But for \( x(s) > \mathcal{M}_{r,t}(t−s) = \frac{1}{2}(x + m_{1/2}(t)) = m_{1/2}(t) + z/2 \), so

\[
u(t−s, x(s)) \leq C_{2}e^{-\sqrt{2}z}.
\]

Hence,

\[
k(t−s, x(s)) \geq 1 − C_{3}e^{−\rho z/\sqrt{2}},
\]

which tends to 1 nicely, as \( z \uparrow \infty \), and so

\[
e^{-2\rho} \exp \left( \int_{0}^{2r} k(t−s, x(s)) ds \right) \geq \exp \left( −2\rho C_{3}e^{−C_{3}} \right).
\]

If \( x − m_{1/2}(t) \geq r \), this tends to one uniformly in \( t > r \), as \( r \uparrow \infty \). Therefore we have the lower bound

\[
\begin{align*}
\quad & \mathcal{M}_{r,t} \quad \text{on} \quad [2r, t−r] \quad \text{(up to constants)}
\end{align*}
\]
\[
    u(t,x) \geq C_1(r)e^{-t} \int_{-\infty}^{\infty} u(r,y) \frac{e^{-(x-y)^2/(2(t-r))}}{\sqrt{2\pi(t-r)}}
    \times P\left[\mathcal{E}_{r,t}^\prime(s) > \mathcal{M}_{r,t}(t-s), \forall s \in [0,t-r]\right] dy,
\]

for \( x \geq m_{1/2}(t) + r \), and \( C_1(r) \uparrow 1 \), as \( r \uparrow \infty \), uniformly in \( t \). This is (6.1.11).

To prove a matching upper bound, one first wants to show that in the Feynman-Kac representation (6.1.2), one may drop the integral from \(-\infty\) to \(-\ln r\). Indeed, by simple change of variables, and using that initial conditions are in \([0,1]\), and \( F(u)/u \) is decreasing in \( u \), we get that

\[
    \int_{-\infty}^{-\ln r} \frac{e^{-(x-y)^2/(2t)}}{\sqrt{2\pi t}} u(0,y) E \left[ \exp \left( \int_0^{t} k(t-s,3x,y(s)) ds \right) \right] dy \leq \int_{-\infty}^{-\ln r} \frac{e^{-(x+y)^2/(2t)}}{\sqrt{2\pi t}} E \left[ \exp \left( \int_0^{t} k(t-s,3x,-y+\ln r(s)) ds \right) \right] dy
\]

Our rough upper (Eq. (6.7.9)) and lower (Eq.4.7.15)) suffice to show that, for \( x \geq m_{1/2}(t) \), uniformly in \( t \),

\[
    \frac{u^H(t,x+\ln r)}{u(t,x)} \downarrow 0, \text{ as } r \uparrow \infty. \tag{6.7.23}
\]

This shows that

\[
    u(t,x) \geq C_2(r) \int_{-\infty}^{\infty} u(0,y)e^{-(x-y)^2/2} E \left[ e^{\int_0^t k(t-s,3x,y(s)) ds} \right] dy, \tag{6.7.24}
\]

where \( C_2(r) \rightarrow 1 \), as \( r \uparrow \infty \), if \( x \geq m_{1/2}(t) \), uniformly in \( t \),

Now we introduce into the Feynman-Kac representation a one in the form

\[
    1 = 1_{\mathcal{E}_{r,t}^\prime(s) > \mathcal{M}_{r,t}(t-s), \forall s \in [0,t]} + 1_{\exists s \in [0,t] : \mathcal{E}_{r,t}^\prime(s) \leq \mathcal{M}_{r,t}(t-s)}. \tag{6.7.25}
\]

Note that by definition of \( \mathcal{M}_r \), the second indicator function is equal to the indicator function of the event \( G_{x,y}(r,t) \). We want to use Proposition 6.40 to show that the contribution coming from the second indicator function is negligible. To be able to do this, we had to make sure that we can drop the contributions from \( y < -\ln r \).

Now we get readily
6.7 Convergence results

\[ \int_{-\infty}^{\infty} \frac{e^{-(x-y)^2/(2t)}}{\sqrt{2\pi t}} u(0,y) \mathbb{E} \left[ \exp \left( \int_0^{t'} k(t-s,3x,y(s)) \,ds \right) \mathbb{1}_{G_{t',y}(x)} \right] \,dy \]

\[ \leq e^{3t} \int_{-\infty}^{\infty} \frac{e^{-(x-y)^2/(2t)}}{\sqrt{2\pi t}} u(0,y) \mathbb{E} \left[ \exp \left( \int_0^{t'-r} k(t-s,3x,y(s)) \,ds \right) \mathbb{1}_{G_{t',y}(x)} \right] \,dy \]

\[ \leq r^{-2} e^t \int_{-\infty}^{\infty} \frac{e^{-(x-y)^2/(2t)}}{\sqrt{2\pi t}} u(0,y) \mathbb{P} \left( G_{t,y}(x,t) \right). \quad (6.7.26) \]

Now Lemma 6.43 provides a bound for the probability in the last line. Inserting this, and observing that we have the lower bound (6.7.15), we see that this is at most a constant times \( u(t,x)/r \), and hence negligible as \( r \uparrow \infty \).

We have arrived at the bound

\[ u(t,x) \leq C_2(r) \int_{-\infty}^{\infty} u(0,y) e^{-(x-y)^2/(2t)} \times \mathbb{E} \left[ \exp \left( \int_0^{t'} k(t-s,3x,y(s)) \,ds \right) \mathbb{1}_{G_{t',y}(x)} \right] \,dy. \quad (6.7.27) \]

Since by definition, \( \mathcal{M} \) imposes no condition on the interval \([0,2r] \), we move the initial condition forward to time \( r \), using that

\[ u(r,y) = \int_{-\infty}^{\infty} u(0,y) e^{-(y-x)^2/(2t)} \times \mathbb{E} \left[ \exp \left( \int_0^{t'} k(r-s,3x,y(s)) \,ds \right) \right] \,dy', \quad (6.7.28) \]

and the Chapman-Kolmogorov equation for the heat kernel, to get from (6.7.27) the bound, for \( x > m_{1/2}(t) \),

\[ u(t,x) \leq C_3(r) e^{t-r} \int_{-\infty}^{\infty} u(r,y) e^{-(x-y)^2/(2(t-r))} \times \mathbb{P} \left[ G_{x,y}(x) > \mathcal{M}_{t}(x,y), \forall s \in [0,t-r] \right] \,dy, \quad (6.7.29) \]

with \( C_3(r) \) tending to one, as \( r \uparrow \infty \). This is (6.1.12).

We have now proven the bounds (6.1.11), and (6.1.12). We still need to prove (6.1.15). Here we need to show the events for the Brownian bridge to stay above \( \mathcal{M} \) is just as unlikely as to stay above \( \mathcal{M} \). These issues were studied in Section 6.5. Let \( \gamma_0 \leq 0 \) be such that \( \int_{\gamma_0} u(0,y) \,dy > 0 \). The bounds we have on \( G_{1/2}(s) \) readily imply that (6.6.30) holds with \( \alpha(t,r) = -\ln(r) \) when \( r \) is large enough. Thus we can use Proposition 6.40 to get that

\[ \frac{\mathbb{P} \left( G_{t,x,y}(s) > \mathcal{M}_{t}(x,y), \forall s \in [0,t-r] \right)}{\mathbb{P} \left( G_{t,x,y}(s) > \mathcal{M}_{t}(x,y), \forall s \in [0,t-r] \right)} = 1 - \varepsilon(r), \quad (6.7.30) \]

for \( t \geq 8r, x - m_{1/2}(t) \geq 8r, \) and \( r \) large enough, with \( \varepsilon(r) \downarrow 0 \), as \( r \uparrow \infty \).
We want to push this to a bound that works for all \( y \). This requires two more pieces of input. The first is a consequence of Proposition 6.44 and Corollary 6.46.

**Corollary 6.47 (Corollary 2 in [22])**. For \( x_2 \geq x_1 \geq 0 \) and \( y_0 \) such that (6.7.1) holds. Then, for \( t \) large enough,

\[
\frac{u(t,x_2)}{u(t,x_1)} \geq C_6 e^{-\frac{(x_2-y_0)^2}{2t}}.
\]

(6.7.31)

Here \( C_6 > 0 \) does not depend on \( t \) or \( x \).

**Proof.** Just combine the upper and lower bound we have by now and use monotonicity properties of the exponential that appear there. \( \Box \)

Next we need the following consequence of (a trivial version of) the FKG inequality.

**Lemma 6.48.** Let \( h, g : \mathbb{R} \rightarrow \mathbb{R} \) be increasing measurable functions. Let \( X \) be a real valued random variable. Then

\[
E[h(X)g(X)] \geq E[h(X)]E[g(X)].
\]

(6.7.32)

**Proof.** Let \( \mu \) denote the law of \( X \). Then

\[
E[h(X)g(X)] - E[h(X)]E[g(X)] = \frac{1}{2} \int \mu(dx) \int \mu(dy) (h(x) - h(y))(g(x) - g(y)).
\]

(6.7.33)

Since both \( f \) and \( g \) are increasing, the two terms in the product have the same sign, and hence the right-hand side of (6.7.33) is non-negative. This implies the assertion of the lemma. \( \Box \)

**Lemma 6.49 (Lemma 8.1 in [22]).** Let \( h : \mathbb{R} \rightarrow [0,1] \) be increasing and let \( g : \mathbb{R} \rightarrow \mathbb{R}_+ \) be such that, for \( x_1 \leq x_2 \), it holds that \( g(x_1) \leq Cg(x_2) \), for some \( C > 0 \). Then, if \( X \) is a real valued random for which \( E[h(X)] \geq 1 - \epsilon \), then

\[
E[g(X)h(X)] \geq (1 - C\epsilon)E[g(X)].
\]

(6.7.34)

**Proof.** Set \( g_1(x) \equiv \sup\{g(y) : y \leq x\} \). Then \( g_1 \) is increasing and positive, so the FKG inequality applies and states that

\[
E[g_1(X)h(X)] \geq E[g_1(X)]E[h(X)] \geq (1 - \epsilon)E[g_1(X)].
\]

(6.7.35)

By assumption, for any \( x \), \( g_1(x) \leq Cg(x) \). Hence

\[
E[g(X)(1-h(X))] \leq CE[g_1(X)(1-h(X))] \leq CE[g_1(X)] - C(1-\epsilon)E[g_1(X)] = C\epsilon E[g_1(X)].
\]

(6.7.36)

Hence
\[ \mathbb{E}[g(X)h(X)] \geq \mathbb{E}[g(X) - C \varepsilon g_1(X)] \geq (1 - \varepsilon)\mathbb{E}[g(X)]. \quad (6.7.37) \]

\[ \square \]

We use Lemma 6.49 in the following construction. Let us denote by \( b(y; x, y_0, t, r) \) the probability density that the Brownian bridge from \( y_0 \) passes through \( y \) at time \( t - r \). This can be expressed in terms of the heat kernel \( p(t, x) \) as

\[ b(y; x, y_0, t, r) = \frac{p(r, y - y_0)p(t - r, x - y)}{p(t, x - y_0)}. \quad (6.7.38) \]

Noting that the condition on the bridge in the denominator of (6.7.30) has bearing only on the time interval \([0, t - r]\), we see that using the Markov property,

\[ \mathbb{P}\left( \dot{z}_{x, y_0}^r(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right) \]
\[ = \int dy b(y; x, y_0, t, r) \mathbb{P}\left( \dot{z}_{x,y}^{t-r}(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right). \quad (6.7.39) \]

We now prepare to use the FKG inequality in the form of Lemma 6.49. For this we represent the left-hand side of (6.7.30) as

\[ \int dy b(y; x, y_0, t, r) \frac{\mathbb{P}\left( \dot{z}_{x,y}^{t-r}(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right)}{\mathbb{P}\left( \dot{z}_{x,y}^r(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right)} \]
\[ = \int \frac{\mathbb{P}\left( \dot{z}_{x,y}^{t-r}(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right)}{\mathbb{P}\left( \dot{z}_{x,y}^r(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right)} \times \frac{\mathbb{P}\left( \dot{z}_{x,y}^r(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right)}{\mathbb{P}\left( \dot{z}_{x,y_0}^r(s) > \mathcal{M}_x^r(t - s), \forall s \in [0, t - r] \right)} b(y; x, y_0, t, r) dy \]
\[ = \int h_{t, r}^y(dy) \mu_{t, r}^y(dy) \quad (6.7.40) \]

Notice that \( \mu_{t, r}^y \) is a probability measure, and \( h_{t, r}^y \) is increasing. Hence, the hypothesis of Lemma 6.49 are satisfied for the expectation of the function \( h_{t, r}^y \) with respect to the probability \( \mu_{t, r}^y \). It remains to choose a function \( g \) such that the expectations in the left- and right-hand sides of (6.7.34) turn into the integrals appearing in (6.1.11) and (6.1.12). For this we chose

\[ g_r(y) \equiv \frac{u(r, y)}{p(r, y - y_0)} \mathbb{1}_{y \geq 0}. \quad (6.7.41) \]

Corollary 6.47 ensures that \( g_r \) satisfies the hypothesis on \( g \) needed in Lemma 6.49 with a constant \( C = 1/C_6 \). With these choices,
\[
\int g_r(y) h_{r,r}^*(y) \mu_{t,r}^*(dy) \tag{6.7.42}
\]
\[
= \frac{\int_0^\infty u(r,y)p(t-r,x-y) \mathbb{P}\left(\mathcal{Y}_{t-r}(s) > \mathcal{M}_{t-r}^r(t-s), \forall s \in [0,t-r]\right) dy}{p(t,x-y_0) \mathbb{P}\left(\mathcal{Y}_{t-r}(s) > \mathcal{M}_{t-r}^r(t-s), \forall s \in [0,t-r]\right)}
\geq (1 - Ce) \frac{\int_0^\infty u(r,y)p(t-r,x-y) \mathbb{P}\left(\mathcal{Y}_{t-r}(s) > \mathcal{M}_{t-r}^r(t-s), \forall s \in [0,t-r]\right) dy}{p(t,x-y_0) \mathbb{P}\left(\mathcal{Y}_{t-r}(s) > \mathcal{M}_{t-r}^r(t-s), \forall s \in [0,t-r]\right)}
\]
This yields
\[
\int_{-\infty}^\infty u(r,y) \frac{e^{-\frac{(x-y)^2}{2(t-r)}}}{\sqrt{2\pi(t-r)}} \mathbb{P}\left(\mathcal{Y}_{t-r}(s) > \mathcal{M}_{t-r}^r(t-s), s \in [0,t-r]\right) dy \tag{6.7.43}
\]
\[
\geq (1 - Ce(r)) \int_0^\infty u(r,y) \frac{e^{-\frac{(x-y)^2}{2(t-r)}}}{\sqrt{2\pi(t-r)}} \mathbb{P}\left(\mathcal{Y}_{t-r}(s) > \mathcal{M}_{t-r}^r(t-s), s \in [0,t-r]\right) dy.
\]
Here we used that the integrals from $-\infty$ to zero are non-negative. Finally, one shows that for the values of $x$ in question, these integrals are also negligible compared to those from $0$ to $\infty$, and hence the integral on the right-hand side can also be replaced by that over the entire real line. This yields the upper bound in (6.1.15). Since the lower bound is trivial, we have proven (6.1.15) and this concludes the proof of Theorem 6.2. \(\square\)

We can now proof the special case of Bramson’s convergence theorem, Theorem 5.9 that is of interest to us.

**Theorem 6.50.** Let $u$ be a solution of the F-KPP equation with $F$ satisfying the standard conditions. Assume that the finite mass condition (6.1.8) holds, and that for some $\eta, M < N > 0$,
\[
\int_x^{x+N} u(0,y) dy > \eta, \quad \text{for} \quad x \leq -M. \tag{6.7.44}
\]
Then
\[
u(t,x + m(t)) \to \mathcal{W}_\gamma(x), \quad \text{as} \quad t \to \infty, \tag{6.7.45}
\]
uniformly in $x \in \mathbb{R}$, where
\[
m(t) = \sqrt{2t} - \frac{3}{2\sqrt{2}} \ln t. \tag{6.7.46}
\]

**Proof.** The proof uses Proposition 6.17. To show that the hypothesis of the proposition are satisfied, we show that, for $x$ large enough,
\[
C e^{-\sqrt{2}\gamma_1^{-1}(x) - \gamma_2(t)} \leq u(t,x + m(t)) \leq C e^{-\sqrt{2}\gamma_1(x) + \gamma_2(t)}, \tag{6.7.47}
\]
where \( \gamma_1(x) \to 1, \) as \( x \uparrow \infty \) and \( \gamma_2(t) \to 0, \) as \( t \uparrow \infty . \) But \( C x e^{-\sqrt{2} x} \) is precisely the tail asymptotics of \( w \sqrt{2} \), so that \( w \sqrt{2} \) can be chosen (remember that we are free to shift the solution to adjust the constant) the hypothesis of Proposition 6.17 are satisfied, and the conclusion of the theorem follows.

Eq. 6.7.47 follows from Theorem 6.2 via the following corollary whose proof we postpone to the next chapter, where various further tail estimates will be proven.

**Corollary 6.51.** Let \( u \) and \( m(t) \) be as in the theorem. Then

\[
\lim_{x \to \infty} \lim_{t \to \infty} e^{x \sqrt{2}} \frac{x}{u(t, x + m(t))} = C, \tag{6.7.48}
\]

for some \( 0 < C < \infty . \)

From here we get the desired control on the tail of \( u \) and hence convergence to the travelling wave solution. \( \Box \)
Chapter 7
The extremal process of BBM

In this chapter we discuss the construction of the extremal process of branching Brownian motion following the paper [6]. This construction has two parts. The first is the proof of convergence of Laplace functionals of the point process of BBM as $t \uparrow \infty$. This will turn out to be doable using estimates on the asymptotics of solutions of the F-KPP equation that are the basis of Bramson’s work. The ensuing representation for Laplace functionals is rather indirect. In the second part, one gives an explicit description of the point process that has this Laplace functional as a cluster point process.

7.1 Limit theorems for solutions

The bounds in (6.1.10) have been used by Chauvin and Rouault to compute the probability of deviations of the maximum of BBM, see Lemma 2 in [28]. We will use their arguments in this slightly more general setting. Our aim is to control limits when $t$ and later $r$ tend to infinity.

We will need to analyse solutions $u(t,x + \sqrt{2}t)$ where $x = x(t)$ depends on $t$ in different ways. First we look at $x$ fixed. This corresponds to particles that are just $\frac{3}{2\sqrt{2}} \ln t$ ahead of the natural front of BBM. We start with the asymptotics of $\psi$.

**Proposition 7.1.** Let the assumptions of Theorem 6.2 be satisfied, and assume in addition that $y_0 \equiv \sup \{ y : u(0,y) > 0 \}$ is finite. Then the following hold for any $z \in \mathbb{R}$:

(i) \[ \lim_{t \uparrow \infty} e^{z\sqrt{2}} \frac{t^{3/2}}{3\sqrt{2}\ln t} \psi(r,t,z + \sqrt{2}r) = \sqrt{\frac{2}{\pi}} \int_0^\infty e^{z\sqrt{2}} u(r,y + \sqrt{2}r) dy = C(r). \]  

(ii) \[ \lim_{r \uparrow \infty} C(r) \equiv C \text{ exists and } C \in (0,\infty). \]  

(7.1.1)
(iii) \[ \lim_{t \to \infty} e^{\frac{y}{\sqrt{2}}} \frac{t^{3/2}}{\frac{3}{2\sqrt{2}}} u(t, z + \sqrt{2}r) = C. \quad (7.1.2) \]

**Proof.** Consider first the limit \( t \uparrow \infty \). For fixed \( z \), the integrand converges pointwise to \( \frac{3}{2\sqrt{2}} y e^{\frac{y}{\sqrt{2}}} u(r, y + \sqrt{2}r) \), as can be seen by elementary calculations. In fact, the integrand in (7.1.1) is given by

\[
\begin{align*}
\int_{2 \pi (t - r) \frac{3}{2\sqrt{2}}}^{t^{3/2}} u(r, y + \sqrt{2}r) e^{\frac{y}{\sqrt{2}}} e^{\frac{(x - y)^2}{2\sqrt{2}}} (1 - e^{-2y \left( r + \frac{3}{2\sqrt{2}} \ln t \right)}) (7.1.3) \\
\sim \int_{2 \pi (t - r) \frac{3}{2\sqrt{2}}}^{t^{3/2}} u(r, y + \sqrt{2}r) e^{\frac{y}{\sqrt{2}}} \sim \frac{1}{\sqrt{2\pi}} u(r, y + \sqrt{2}r) e^{\frac{y}{\sqrt{2}}}.
\end{align*}
\]

To deduce convergence of the integral, we need to show that we can use dominated convergence. We see from the first line in (7.1.3) that the integrand in (6.1.9) after multiplication with the prefactor in (7.1.1) is bounded from above by

\[ \text{const.} y e^{\frac{y}{\sqrt{2}}} u(r, y + \sqrt{2}r). \quad (7.1.4) \]

Thus we need an a priori bound on \( u \). By the Feyman-Kac representation and the fact that \( u(0, y) = 0 \), for \( y > y_0 \), implies that

\[ u(r, x) \leq e^r \int_{-\infty}^{y_0} \frac{1}{\sqrt{2\pi r}} e^{-\frac{(y-x)^2}{2r}} dy \leq \begin{cases} C, & x \leq y_0, \\ C e^{-\frac{(y-x)^2}{2r}}, & x \geq y_0. \end{cases} \quad (7.1.5) \]

Setting \( x = y + \sqrt{2}r \) and inserting this into (7.1.4), we can bound the integrand in (6.1.9) by (ignoring constants)

\[ ye^{\frac{y}{\sqrt{2}}} 1_{x \leq y_0 - \sqrt{2}r} + ye^{\frac{y}{\sqrt{2}}} e^{-\frac{(y + \sqrt{2}r) - y_0)^2}{2r}} 1_{y \geq y_0 - \sqrt{2}r} \]

\[ \leq C(r, y_0) 1_{y \leq y_0 - \sqrt{2}r} + C(r, y_0) ye^{-\frac{(y + y_0)^2}{2r}} 1_{y \geq y_0 - \sqrt{2}r} \]

which is integrable in \( y \) uniformly in \( t \). Hence Lebesgue’s dominated convergence theorem applies and the first part of the proposition follows.

The more interesting task is to show now that \( C(r) \) converges to a non-trivial limit, as \( r \uparrow \infty \). By Theorem 6.2, for \( r \) large enough,

\[ \lim_{r \uparrow \infty} e^{\frac{y}{\sqrt{2}}} \frac{t^{3/2}}{\frac{3}{2\sqrt{2}}} u(r, z + \sqrt{2}t) \leq \gamma(r) \lim_{r \uparrow \infty} e^{\frac{y}{\sqrt{2}}} \frac{t^{3/2}}{\ln t} \psi(r, t, z + \sqrt{2}t) = C(r) \gamma(r), \quad (7.1.7) \]

and
7.1 Limit theorems for solutions

\[ \liminf_{t \to \infty} e^{\sqrt{2} t^{3/2}} \frac{1}{3\sqrt{2} \ln t} u(t, z + \sqrt{2} t) \geq \gamma(r)^{-1} \lim_{t \to \infty} e^{\sqrt{2} t^{3/2}} \frac{1}{\ln t} \psi(r, t, z + \sqrt{2} t) = C(r) \gamma(r)^{-1}, \]

(7.1.8)

These bounds hold for all \( r \), and since \( \gamma(r) \to 1 \), we can conclude that the left-hand sides is bounded from above by \( \liminf_{r \to \infty} C(r) \) and from below by and \( \limsup_{r \to \infty} C(r) \). So the \( \limsup \) \( \leq \liminf \) and hence \( \lim_{r \to \infty} C(r) = C \) exists. As a byproduct, we see that also the left-hand sides of (7.1.7) and (7.1.8) hold for any \( r \). But for large enough finite \( r \), the constants \( C(r) \) are strictly positive and finite by the representation (7.1.1), and this shows that \( C \) is strictly positive and finite. \( \square \)

We will make a slight stop on the way and show how the sharp asymptotics for the upper tail of solutions follows from what we just did.

**Corollary 7.2.** Let \( u \) be as in the preceding proposition. Then

\[ \lim_{z \to \infty} \lim_{t \to \infty} e^{\sqrt{2} t^{3/2}} \frac{1}{z} u(t, z + m(t)) = C, \]

(7.1.9)

where \( 0 < C < \infty \) is the same constant as in Proposition 7.1

**Proof.** Note that

\[ \psi(r, t, z + m(t)) = \psi(r, t, t + \frac{1}{2\sqrt{2}} \ln t). \]

(7.1.10)

Then, by the same arguments as in the proof of Proposition 7.1, we get that

\[ \lim_{t \to \infty} e^{\sqrt{2} t^{3/2}} \frac{1}{z} \psi(r, t, z + \sqrt{2} t) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} ye^{\sqrt{2} y} u(r, y + \sqrt{2} r) dy = C(r). \]

(7.1.11)

Here \( C(r) \) the same constant as before. Now we can choose \( z > 8r \) in order to be able to apply Theorem 6.2 and then let \( r \to \infty \). This yields (7.1.9). \( \square \)

The next lemma is a variant of the previous proposition for the case when \( x \sim \sqrt{t} \).

**Lemma 7.3.** Let \( u \) be a solution to the F-KPP equation (5.4.2) with initial condition satisfying the assumptions of Theorem 5.9 and

\[ y_0 \equiv \sup\{y : u(0, y) > 0\} < \infty. \]

(7.1.12)

Then, for \( a > 0 \), and \( y \in \mathbb{R} \),

\[ \lim_{t \to \infty} \frac{e^{\sqrt{2} a \sqrt{t}/3}}{a \sqrt{t}} \psi(r, t, y + a \sqrt{t} + \sqrt{2} t) = C(r) e^{-\sqrt{2} y} e^{-a^2/2} \]

(7.1.13)

where \( C(r) \) is the constant from above. Moreover, the convergence is uniform for \( a \) in compact sets.
Proof. The structure of the proof is the same as in the proposition. Compute the pointwise limit and produce an integrable majorant from a bound using the linearised F-KPP equation. □

It will be very important to know that as $r \uparrow \infty$, in the integral representing $C(r)$, only the $y$’s that are of order $\sqrt{r}$ give a non-vanishing contribution. The precise version of this statement is the following lemma.

**Lemma 7.4.** Let $u$ be a solution to the F-KPP equation (5.4.2) with initial condition satisfying the assumptions of Theorem 5.9 and

$$y_0 \equiv \sup\{y : u(0, y) > 0\} < \infty.$$ (7.1.14)

Then for any $z \in \mathbb{R}$:

$$\lim_{r \downarrow 0} \lim_{A_1 \uparrow \infty} \limsup_{r \uparrow \infty} \int_0^{A_1 \sqrt{r}} u(r, z + y + \sqrt{2r}) ye^{y\sqrt{2}} dy = 0$$ (7.1.15)

$$\lim_{r \uparrow \infty} \limsup_{A_2 \uparrow \infty} \limsup_{r \uparrow \infty} \int_{A_2 \sqrt{r}}^{\infty} u(r, z + y + \sqrt{2r}) ye^{y\sqrt{2}} dy = 0.$$ (7.1.16)

Proof. Again it is clear that we need some a priori information on the behaviour of $u$. This time this will be obtained by comparing to a solution of the F-KPP equation with Heavyside initial conditions, which we know are probabilities for the maximum of BBM. Namely, by assumption, $u(0, y) \leq 1_{\{y < y_0 + 1\}}$. It follows from the representation of $u$ as an expectation (5.4.1) the $u$ is monotone in the initial conditions and thus

$$u(r, z + y + \sqrt{2r}) \leq P\left(\max_{k \leq n(r)} x_k(r) - \sqrt{2r} > y_0 + 1 + z + y\right).$$ (7.1.17)

For the probabilities we have a sharp estimate given in Lemma 6.45 in the previous chapter.

We use this lemma with $z \sim \sqrt{t}$, so the $z^2/t$ term is relevant. The last term involving $\ln t/t$ can, however, be safely ignored. The lemma follows in a rather straightforward and purely computational way once we insert this bound in integrals. The details can be found in [6]. Note that the precise bound (6.7.13) is needed and the trivial bound comparing to iid variables would not suffice. □

The following lemma shows how $C = \lim_{r \uparrow \infty} \sqrt{\frac{2}{\pi}} \int_0^{\infty} u(r, y + \sqrt{2r}) ye^{y\sqrt{2}} dy$ behaves when the spatial argument of $u$ is shifted.

**Lemma 7.5.** Let $u$ be a solution to the F-KPP equation (5.4.2) with initial condition satisfying the assumptions of Theorem 5.9 and

$$y_0 \equiv \sup\{y : u(0, y) > 0\} < \infty.$$ (7.1.17)

Then for any $z \in \mathbb{R}$:
7.1 Limit theorems for solutions

\[
\lim_{r \to \infty} \sqrt{\frac{t}{\pi}} \int_0^\infty ye^{\sqrt{2}y} u(r, z + y + \sqrt{2}r) dy = e^{-\sqrt{2}z} \lim_{r \to \infty} \sqrt{\frac{t}{\pi}} \int_0^\infty ye^{\sqrt{2}y} u(r, y + \sqrt{2}r) dy = Ce^{-\sqrt{2}z}.
\]

**Proof.** The proof is obvious from the fact proven in Lemma 7.4. Namely, changing variables,

\[
\int_0^\infty ye^{\sqrt{2}y} u(r, z + y + \sqrt{2}r) dy = e^{-\sqrt{2}z} \int_0^\infty (y - z) e^{\sqrt{2}y} u(r, y + \sqrt{2}r) dy
\]

Now the part of the integral from \( z \) to zero cannot contribute, as it does not involve \( y \) of the order of \( \sqrt{r} \), and for the same reason replacing \( y \) by \( y - z \) changes nothing in the limit. □

The following is a slight generalisation of Lemma 7.5:

**Lemma 7.6.** Let \( h(x) \) be continuous function that is bounded and is zero for \( x \) small enough. Then

\[
\lim_{t \to \infty} \sqrt{\frac{2}{\pi}} \int_{-\infty}^0 \mathbb{E} \left[ h(y + \max_{i \leq t} x_i(t) - \sqrt{2}t) \right] (-y) e^{-\sqrt{2}y} dy = \int_{-\infty}^0 h(z) \sqrt{2} Ce^{-\sqrt{2}z} dz,
\]

where \( C \) is the constant appearing in the in the law of the maximum, Theorem 5.16

**Proof.** Consider first the case when \( h(x) = \mathbb{1}_{(b,\infty)}(x) \), for \( b \in \mathbb{R} \).

\[
\lim_{t \to \infty} \sqrt{\frac{2}{\pi}} \int_{-\infty}^0 \mathbb{P} \left[ y + \max_{i \leq t} x_i(t) - \sqrt{2}t > b \right] (-y) e^{-\sqrt{2}y} dy
\]

\[
= \lim_{t \to \infty} \int_{0}^0 \mathbb{P} \left[ \max_{i \leq t} x_i(t) > b + y + \sqrt{2}t \right] \sqrt{\frac{2}{\pi}} y e^{\sqrt{2}y} dy
\]

\[
= \lim_{t \to \infty} \int_{0}^0 u(t, b + y + \sqrt{2}t) \sqrt{\frac{2}{\pi}} y e^{\sqrt{2}y} dy,
\]

where \( u \) solves the F-KPP equation (5.4.2) with initial condition \( u(0, x) = \mathbb{1}_{x<0} \). Hence the right-hand side of (7.1.21) converges, by Lemma 7.5 to

\[
Ce^{-\sqrt{2}b} = C \int_{b}^\infty e^{-\sqrt{2}z} \sqrt{2} dz = C \int_{-\infty}^\infty h(z) e^{-\sqrt{2}z} \sqrt{2} dz.
\]

To pass to the general case (7.1.20), note that by linearity, (7.1.22) holds for linear combinations of indicator functions of intervals. The general case when follows from a monotone class type argument. □
7.2 Existence of a limiting process

We now look at the construction of the limit of the point processes

$$\mathcal{E}_t \equiv \sum_{k \leq n(t)} \delta_{x_k(t) - m(t)}.$$  \hfill (7.2.1)

7.2.1 Convergence of the Laplace functional

To prove the existence of the limiting extremal process we show tightness and the convergence of a sufficiently rich class of Laplace functionals

$$\psi_t(\phi) \equiv \mathbb{E}\left[ \exp \left( - \int \phi(y) \mathcal{E}_t(dy) \right) \right].$$  \hfill (7.2.2)

From Lemma 5.5 one has that the Laplace functional related to a solution $u$ of the F-KPP equation with initial conditions $u(0,x) = 1 - e^{-\phi(-x)}$ via

$$1 - \psi_t(\phi) = u(t,m(t)).$$  \hfill (7.2.3)

As we want to use Bramson’s main theorem, Theorem 5.9, we use functions $\phi$ such that the hypotheses (i) and (ii) in that theorem are satisfied. A convenient choice is the class of functions of the form

$$\phi(x) = \sum_{\ell=1}^k c_{\ell} \mathbb{1}_{x > u_{\ell}},$$  \hfill (7.2.4)

with $c_{\ell} > 0$ and $u_{\ell} \in \mathbb{R}$. Set $g(x) \equiv 1 - e^{-\phi(-x)}$. Clearly $g(x)$ vanishes for $x > -\min_{\ell}(u_{\ell})$, while for $x < -\max_{\ell}(u_{\ell})$, we have that $g(x) \geq 1 - e^{-\sum_{\ell=1}^k u_{\ell}} > 0$. This implies that $g(x)$ satisfies both hypotheses for the initial conditions of Theorem 5.9. On the other hand, if $u(t,x)$ solves the F-KPP equation with initial condition $u(0,x) = g(x)$,

$$1 - u(t,x) = \mathbb{E}\left[ \prod_{i=1}^{n(t)} e^{-\sum_{\ell=1}^k \phi(x_i(t)-x)} \right]$$  \hfill (7.2.5)

is the Laplace functional for our choice of $\phi$. Thus, if we can show that $u(t,x+m(t))$ converges to a non-trivial limit, we obtain the following theorem.

**Theorem 7.7.** The point process $\mathcal{E}_t = \sum_{k \leq n(t)} \delta_{x_k(t) - m(t)}$ converges in law to a point process $\mathcal{E}$.

**Proof.** First, the sequence of processes $\mathcal{E}_t$ is tight by Corollary 2.21 since that for any bounded interval, $B \subset \mathbb{R}$,
7.2 Existence of a limiting process

\[ \lim_{N \to \infty} \lim_{t \to \infty} \mathbb{P}(\sigma_t(B) \geq N) = 0, \quad (7.2.6) \]

so the limiting point process must be locally finite. (Basically, if (7.2.6) failed then the maximum of BBM would have to be much larger than it is know to be). We now show (7.2.6). Clearly it is enough to show this for \( B = [y, \infty) \). Assume that (7.2.6) does not hold. Then there exists \( \varepsilon > 0 \) and a sequence \( t_N \uparrow \infty \) there exists such that for all \( N < \infty \),

\[ \mathbb{P}(\sigma_{t_N}([y, \infty)) \geq N) \geq \varepsilon. \quad (7.2.7) \]

On the other hand, we know that for any \( \delta > 0 \), we can find \( a_\delta \in \mathbb{R} \), such that

\[ \lim_{t \to \infty} \mathbb{P}\left( \max_{k \leq n(t)} x_k(t) \leq m(t) + a_\delta \right) \geq 1 - \delta. \quad (7.2.8) \]

Now chose \( \delta = \varepsilon / 2 \). It follows that, for \( N \) large enough

\[ \mathbb{P}\left( \{\sigma_t(B) \geq N\} \cap \{\max_{k \leq n(t) + 1} x_k(tN + 1) \leq m(tN + 1) + a_{\varepsilon/2}\} \right) \geq \varepsilon / 2. \quad (7.2.9) \]

But this probability is bounded from above by the probability that the offspring of the \( N \) particles at time \( t_N \) that must be above \( m(t_N) + y \) remain below \( m(t_N + 1) + a_{\varepsilon/2} \), which is smaller than

\[ \prod_{i=1}^{N} \mathbb{P}\left( x_i(t_N) + \max_{k \leq n_i(1)} x_k^{(j)}(1) \leq m(t_N + 1) + a_{\varepsilon/2}[x_j(t_N) \geq m(t_N) + y] \right) \leq \left[ \mathbb{P}\left( \max_{k \leq n(1)} x_k(1) \leq m(t_N + 1) - m(t_N) + a_{\varepsilon/2} - y \right) \right]^N \quad (7.2.10) \]

But the probability in the last line is certainly smaller than 1, and hence choosing \( N \) large enough, the expression in the last line is strictly smaller than \( \varepsilon / 2 \), which leads to a contradiction. This proves (7.2.6).

For any \( \phi \) of the form (7.2.4) we now know from Theorem 5.9 that \( u(t, x + m(t)) \) converges, as \( t \to \infty \), to a travelling wave \( w_{\sqrt{2}} \). Hence

\[ \lim_{t \to \infty} \psi_t(\phi) = \psi(\phi) \quad (7.2.11) \]

exists, and is strictly smaller than one. But
\[
\psi_t(\phi) = E \left[ \exp \left( - \int \phi \, dE_t \right) \right] \\
= E \left[ \prod_{k \leq n(t)} \exp \left( - \phi(x_k(t) - m(t)) \right) \right] \\
= E \left[ \prod_{k \leq n(t)} g(m(t) - x_k(t)) \right] = u(t, 0 + m(t)),
\]
and therefore \( \lim_{t \to \infty} \psi_t(\phi) \equiv \psi(\phi) \) exists which proves (7.2.11). \( \square \)

**Remark 7.8.** The particular choice of the functions \( \phi \) for which we prove convergence of the Laplace functionals is made since they satisfy the hypothesis of Theorem 5.9. It implies, of course, convergence for all \( \phi \in C^+_{c}(\mathbb{R}) \) (and more). They do not satisfy the assumption (ii) in Bramson’s theorem, so convergence of the corresponding solutions cannot be uniform in \( x \), but that is not important. In [6] convergence was proven directly for \( \phi \in C^+_{c}(\mathbb{R}) \) using a truncation procedure. Essentially, it makes no difference whether the support of \( \phi \) is bounded from above or not, since the maximum of the point process \( E_t \) is finite almost surely, uniformly in \( t \).

### 7.2.2 Flash back to the derivative martingale.

Having established the tail asymptotics for solutions of the F-KPP equation, one might be tempted to compute the law of the maximum (or Laplace functionals) by recursion at an intermediate time \( r \), e.g.

\[
E \left[ \exp \left( - \int \phi(-x + y) \, dE_t(dy) \right) \right] = E \left[ \exp \left( - \int \phi(-x + y) \, dE_t(dy) \right) \bigg| \mathcal{F}_r \right] \\
= E \left[ \prod_{i=1}^{n(r)} \exp \left( - \int \phi(-x + x_i(r) - m(t) + m(t-r) + y) \, dE_t^{(i)}(dy) \right) \bigg| \mathcal{F}_r \right] \\
\approx E \left[ \prod_{i=1}^{n(r)} \left( 1 - \nu(t-r, -x + x_i(r) - \sqrt{2r}) \right) \bigg| \mathcal{F}_r \right], \tag{7.2.12}
\]

where \( \mathcal{F}^{(i)}_t \) are iid copies on BBM and

\[
\nu(s, x - x_i(r) + \sqrt{2r} + m(s)) = 1 - \exp \left( - \int \phi(-x + x_i(r) - \sqrt{2r} + y) \, dE_t^{(i)}(dy) \right).
\tag{7.2.13}
\]

We have used that \( m(t) - m(t-r) = \sqrt{2r} + O(r/t) \). We want to use the fact (see Theorem 5.2) that the particles that will show up in the support of \( \phi \) at time \( t \) must come from particles in an interval \( (\sqrt{2r} - c_1 \sqrt{r}, \sqrt{2r} - c_2 \sqrt{r}) \), if \( r \) is large enough and \( c_1 > c_2 > 0 \) become large. Hence we may assume that the conditional expectations that appear in (7.2.12) can be replaced by their tail asymptotics, i.e.
7.2 Existence of a limiting process

\[ v(t-r,x-x_i(r)+\sqrt{2}r+m(t-r)) \sim C(x-x_i(r)+\sqrt{2}r)e^{-\sqrt{2}(x-x_i(r)+\sqrt{2}r)}. \] (7.2.14)

Hence we would get,

\[ \lim_{t \to \infty} E \left[ \exp \left( -\int \phi(x+y) \mathcal{E}_t(dy) \right) \right] = \lim_{r \to \infty} E \left[ \exp \left( -C \sum_{i=1}^{n(r)} (x-x_i(r)+\sqrt{2}r)e^{-\sqrt{2}(x-x_i(r)+\sqrt{2}r)} \right) \right] = \lim_{r \to \infty} E \left[ \exp \left( -CZ_t e^{-\sqrt{2}x} - CY_t x e^{-\sqrt{2}x} \right) \right]. \] (7.2.15)

with \( Z_t \) and \( Y_t \) the martingales from Section 2.4. The constant \( C \) will of course depend in general on the initial data for \( \nu \), i.e. of \( \phi \).

Now one could argue like Lalley and Sellke that \( Y_t \) must converge to a non-negative limit, implying that \( \sqrt{2}r-x_i(r) \uparrow \infty \) if this limit is strictly positive. Then again this implies that \( Y_t \to 0 \), a.s., and hence

\[ \lim_{t \to \infty} E \left[ \exp \left( -\int \phi(x+y) \mathcal{E}_t(dy) \right) \right] = \lim_{r \to \infty} E \left[ e^{-CZ_t e^{-\sqrt{2}x}} \right]. \] (7.2.16)

The argument above is not quite complete. In the next subsection we will show that the final result is nonetheless true.

7.2.3 A representation for the Laplace functional

In Subsection 7.2 we have shown convergence of the Laplace functional. The following proposition exhibits the general form of the Laplace functional of the limiting process. In this section we will always assume that \( \phi \) is a function of the form (7.2.4).

The following proposition exhibits the general form of the Laplace functional of the limiting process.

**Proposition 7.9.** Let \( \mathcal{E}_t \) be the process (7.2.1). For \( \phi \) be as given in (7.2.4) and any \( x \in \mathbb{R} \),

\[ \lim_{t \to \infty} E \left[ \exp \left( -\int \phi(y+x) \mathcal{E}_t(dy) \right) \right] = E \left[ \exp \left( -C(\phi)Z_t e^{-\sqrt{2}x} \right) \right] \] (7.2.17)

where, for \( u(t,y) \) the solution of the F-KPP equation with initial condition \( u(0,y) = 1 - e^{-\phi(-y)} \),

\[ C(\phi) = \lim_{t \to \infty} \sqrt{\frac{2}{\pi}} \int_0^\infty u(t,y+\sqrt{2}r)ye^{\sqrt{2}r}dy \] (7.2.18)

is a strictly positive constant depending on \( \phi \) only, and \( Z_t \) is the derivative martingale.
Proof. The initial conditions \( u(0, y) \) satisfy the hypothesis that were required in the previous results on sharp approximation. The existence of the constant \( C(\phi) \) then follows from Proposition 7.1.

On the other hand, it follows from Theorem 5.9 and the Lalley-Sellke representation, that

\[
\lim_{t \to \infty} u(t, x + m(t)) = 1 - \mathbb{E} \left[ \exp \left( - C(\phi) Z e^{-\sqrt{2} x} \right) \right],
\tag{7.2.19}
\]

for some constant \( C(\phi) \) (recall that the relevant solution of (5.5.2) is unique up to shifts (see Lemma 5.8), which translates into uniqueness of the representation (7.2.17) up to the choice of the constant \( C \)). Clearly, form the asymptotic of solutions, we know that

\[
\lim_{x \to \infty} \frac{1 - \mathbb{E} \left[ \exp \left( - C Z e^{-\sqrt{2} x} \right) \right]}{x e^{-\sqrt{2} x}} = C.
\tag{7.2.20}
\]

Thus all what is left is to identify this constant with \( C(\phi) \). But from Corollary 7.2 it follows that \( C \) must be \( C(\phi) \). This is essentially done by rerunning Proposition (7.1) with \( \sqrt{2} t \) replaced by the more precise \( m(t) \).

This proves (7.2.17). \( \square \)

7.3 Interpretation as cluster point process

In the preceding section we have given a full construction of the Laplace functional of the limiting extremal process of BBM and have given a representation of it in Proposition 7.9. Note that all the information on the limiting process is contained in the way how the constant \( C(\phi) \) depends on the function \( \phi \). The characterisation of this dependence via a solution of the F-KPP equation with initial condition given in terms of \( \phi \) does not appear very revealing at first sight. In the following we will remedy this by giving explicit probabilistic descriptions of the underlying point process.

7.3.1 Interpretation via an auxiliary process

We will construct an auxiliary class of point processes that a priori have nothing to do with the real process of BBM. Let \( (\eta_i; i \in \mathbb{N}) \) denote the atoms of a Poisson point process on \((-\infty, 0)\) with intensity measure

\[
\sqrt{\frac{2}{\pi}} (-x) e^{-\sqrt{2} x} dx.
\tag{7.3.1}
\]

For each \( i \in \mathbb{N} \), consider independent BBM’s \( \{X_k^{(i)}(t), k \leq n(i)(t)\} \). Note that, for each \( i \in \mathbb{N} \),
This follows, e.g., from the fact that the martingale $Y(t)$ defined in (5.6.7) converges to zero, a.s.. The auxiliary point process of interest is constructed from these ingredients as

$$
\Pi_t \equiv \sum_i \sum_{k=1}^{n_i(t)} \delta_{\frac{\ln Z + \eta_i x^{(i)}(t) - \sqrt{2t}}{\sqrt{2}}}
$$

where $Z$ has the same law as limit of the derivative martingale. The existence and non-triviality of the process in the limit $t \uparrow \infty$ is not obvious, especially in view of (7.3.2). We will show that not only it exists, but it has the same law as the limit of the extremal process of BBM.

**Theorem 7.10.** Let $\delta_t$ be the extremal process (7.2.1) of BBM. Then

$$
\lim_{t \uparrow \infty} \delta_t \stackrel{\text{law}}{=} \lim_{t \uparrow \infty} \Pi_t.
$$

**Proof.** The proof of this result just requires the computation of the Laplace transform, which we are already quite skilled in.

The Laplace functional of $\Pi_t$ using the form of the Laplace functional of a Poisson process reads

$$
E \left[ \exp \left( - \int \phi(x) \Pi_t(dx) \right) \right] = E \left[ \exp \left( - \sum_i \left( \sum_{k=1}^{n_i(t)} \phi \left( \eta_i + \ln Z + \frac{x^{(i)}(t) - \sqrt{2t}}{\sqrt{2}} \right) \right) \right) \right] = E \left[ \exp \left( - \sum_i \Theta_i^{(i)}(\eta_i) \right) \right],
$$

where we set

$$
\Theta_i^{(i)}(x) \equiv \sum_{k=1}^{n_i(t)} \phi \left( x + \ln Z + \frac{x^{(i)}(t) - \sqrt{2t}}{\sqrt{2}} \right).
$$

Now we want to compute the expectation. Using the independence of the Cox process and the processes $x^{(i)}(t)$, and using the explicit form of the Laplace functional of Poisson processes, we see that, (7.3.5) is equal to

$$
E \left[ \exp \left( \mathbb{E} \left[ \left( e^{-\Theta^{(i)}(x)} - 1 \right) \bigg| \sigma(Z) \right] \right) \right] = E \left[ \exp \left( \int_{-\infty}^{0} \mathbb{E} \left[ \left( e^{-\Theta^{(i)}(x)} - 1 \right) \bigg| \sigma(Z) \right] \sqrt{\frac{2}{\pi}}(-x)e^{-\sqrt{2x}}dx \right) \right].
$$

Note that the conditional expectation in the exponent is just the expectation with respect to one BBM, $x^{(1)}$. The outer expectation is just the average with respect to
the derivative martingale $Z$. Now the exponent is explicitly given as

$$\int_{-\infty}^{0} \left( \mathbb{E} \prod_{k=1}^{n(t)} \left( e^{-\phi(x+Z+\sqrt{2}x(t)-\sqrt{2}Z)} \right) - 1 \right) \sqrt{\frac{2}{\pi}} e^{-\sqrt{2}x} dx$$  \hspace{1cm} (7.3.8)

$$= - \int_{-\infty}^{0} u(t, -x + \sqrt{2}t - Z/\sqrt{2}) \sqrt{\frac{2}{\pi}} e^{-\sqrt{2}x} dx$$

$$= - \int_{0}^{\infty} u(t, x + \sqrt{2}t - Z/\sqrt{2}) \sqrt{\frac{2}{\pi}} xe^{\sqrt{2}x} dx,$$

where $u$ is a solution of the F-KPP equation with initial condition $1 - e^{-\phi(-x)}$. Hence, by (7.1.18),

$$\lim_{t \uparrow \infty} \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} u(t, x + \sqrt{2}t - \frac{1}{\sqrt{2}} \ln Z) xe^{\sqrt{2}x} dx$$  \hspace{1cm} (7.3.9)

$$= Z \sqrt{\frac{2}{\pi}} \lim_{t \uparrow \infty} \int_{0}^{\infty} u(t, x + \sqrt{2}t) xe^{\sqrt{2}x} dx = ZC(\phi),$$

where the last equality follows by Proposition 7.9. Thus we arrive finally at the

$$\lim_{t \uparrow \infty} \mathbb{E} \left[ \exp \left( - \int \phi(x) \Pi_t(dx) \right) \right] = \mathbb{E} \left[ e^{-ZC(\phi)} \right].$$

This implies that the Laplace functionals of $\lim_{t \uparrow \infty} \Pi_t$ and of the extremal process of BBM are equal, and proves the proposition. \qed

### 7.3.2 The Poisson process of cluster extremes

We will now give another interpretation of the extremal process. In the paper [4], the following result was shown: consider the ordered enumeration of the particles of BBM at time $t$,

$$x_1(t) \geq x_2(t) \geq \cdots \geq x_{n(t)}(t).$$  \hspace{1cm} (7.3.10)

Fix $r > 0$ and construct the equivalence classes of these particles such that each class consists of particles that have a common ancestor at time $s > t - r$. For each of these classes, we can select as a representative its largest member. This can be constructed recursively as follows:

$$i_1 = 1,$$

$$i_k = \min \{ j > i_{k-1} : q(i_i, j) < t - r, \forall \ell < k - 1 \}.$$  \hspace{1cm} (7.3.11)

This is repeated to exhaustion and provides the desired decomposition. We denote the resulting number of classes by $n^*(t)$. We can think of the $x_{i_k}(t)$ as the heads of families at time $t$. We then construct the point processes
7.3 Interpretation as cluster point process

\[ \Theta'_r \equiv \sum_{k=1}^{n^*(t)} \delta_{x_{ik}(t) - m(t)}. \]  

(7.3.12)

In [4] the following result was proven.

**Theorem 7.11.** Let \( \Theta'_r \) be defined above. Then

\[ \lim_{r \uparrow \infty} \lim_{t \uparrow \infty} \Theta'_r = \Theta, \]

(7.3.13)

where convergence is in law and \( \Theta \) is the random Poisson process (Cox process) with intensity measure \( CZ\sqrt{2}e^{-\sqrt{2}x}dx \), \( Z \) being the derivative martingale, and \( C \) the constant from Theorem 5.16.

**Remark 7.12.** The thinning out of the extremal process is removing enough correlation to Poissonise the process. This is rather common in the theory of extremal processes, see e.g. [9, 10].

I will not prove this theorem here. Rather, I will show how this result links up to the representation of the extremal process of BBM given above.

**Proposition 7.13.** Let

\[ \Pi'_{\text{ext}} t \equiv \sum_i \delta_{\sqrt{2} \ln Z + \eta_i + M_i(t) - \sqrt{2}t} \]

(7.3.14)

where \( M^{(i)}(t) \equiv \max_k x_k^{(i)}(t) \), i.e. the point process obtained by retaining from \( \Pi \) the maximal particles of each of the BBM’s. Then

\[ \lim_{t \uparrow \infty} \Pi'_{\text{ext}} t \overset{\text{law}}{=} \text{PPP} \left( Z\sqrt{2}Ce^{-\sqrt{2}x}dx \right) \]

(7.3.15)

as a point process on \( \mathbb{R} \), where \( C \) is the same constant appearing the law of the maximum. In particular, the maximum of \( \lim_{t \uparrow \infty} \Pi'_{\text{ext}} t \) has the same law as the limit law of the maximum of BBM.

**Proof.** Consider the Laplace functional of our thinned auxiliary process,

\[ \mathbb{E} \left[ \exp \left( - \sum_i \phi(\eta_i + M^{(i)}(t) - \sqrt{2}t) \right) \right] \]

(7.3.16)

Since the \( M^{(i)} \)'s are i.i.d., and denoting by \( \mathcal{F}_{\eta} \) the \( \sigma \)-algebra generated by the Poisson point process of the \( \eta \)'s, we get that
\[ E \left[ \exp \left( -\sum_i \phi(\eta_i + M_i^0(t) - \sqrt{2t}) \right) \right] \quad (7.3.17) \]

\[ = E \left[ \prod_i E \left( \exp \left( -\phi(\eta_i + M_i^0(t) - \sqrt{2t}) \right) \right) \right] \]

\[ = E \left[ \exp \left( -\sum_i g(\eta_i) \right) \right], \]

where

\[ g_i^{(i)}(z) \equiv \phi(z + M(t) - \sqrt{2t}). \quad (7.3.18) \]

Proceeding as in the proof of Theorem 7.10,

\[ E \left[ \exp \left( -\sum_i g(\eta_i) \right) \right] \quad (7.3.19) \]

\[ = E \left[ \exp \left( -\int_{-\infty}^{0} E_{M_t} \left[ e^{-g(y + \ln Z/\sqrt{2} + M(t) - \sqrt{2t})} - 1 \right] \sqrt{\frac{2}{\pi}} (y) e^{-\sqrt{2}y} dy \right) \right], \]

where \( E_{M_t} \) denote expectation w.r.t. the max of BBM at time \( t \). Using Lemma 7.5 with \( h(x) = 1 - e^{-\phi(-x)} \), we see that

\[ \lim_{t \to \infty} \int_{-\infty}^{0} E_{M_t} \left[ 1 - e^{-\phi(y + \ln Z/\sqrt{2} + M(t) - \sqrt{2t})} \right] \sqrt{\frac{2}{\pi}} (y) e^{-\sqrt{2}y} dy \quad (7.3.20) \]

\[ = \int_{-\infty}^{\infty} (1 - e^{-\phi(a)}) C \sqrt{2} e^{-\sqrt{2}a} da, \]

which is the Laplace functional of the PPP with intensity \( ZC e^{-\sqrt{2}a} da \) on \( R \). The assertion of Proposition 7.13 follows. \( \square \)

So this is nice. The process of the maxima of the BBM’s in the auxiliary process is the same Poisson process as the limiting process of the heads of families. This gives a clear interpretation of the BBM’s in the auxiliary process.

### 7.3.3 More on the law of the clusters

Let us continue to look at the auxiliary process. We know that it is very hard for any of the \( M_i^0(t) \) to exceed \( \sqrt{2t} \) and thus to end up on above a level \( a \) as \( t \to fty \). Therefore, many of the atoms \( \eta_i \) of the Poisson process will not have relevant offspring.

The following proposition states that there is a narrow window deep below zero from which all relevant particles come. This is analogous to the observation in Lemma 7.1.14.
Proposition 7.14. For any \( y \in \mathbb{R} \) and \( \varepsilon > 0 \) there exist \( 0 < A_1 < A_2 < \infty \) and \( t_0 \) depending only on \( y \) and \( \varepsilon \), such that

\[
\sup_{t \geq 0} \mathbb{P} \left( \exists_{i,k} : \eta_i + x_k^{(i)}(t) - \sqrt{2}t \geq z, \wedge \eta_i \notin [-A_1 \sqrt{t}, -A_2 \sqrt{t}] \right) < \varepsilon. \tag{7.3.21}
\]

Proof. Throughout the proof, the probabilities are considered conditional on \( Z \). Clearly we have

\[
\mathbb{P} \left( \exists_{i,k} : \eta_i + x_k^{(i)}(t) - \sqrt{2}t \geq z, \text{ but } \eta_i \geq -A_1 \sqrt{t} \right) \leq \sum_{i,j} \mathbb{E} \left[ \mathbb{P} \left( M^{(i)}(t) - \sqrt{2}t + \eta_i \geq y | \eta_i \right) C(-y)e^{-\sqrt{2}y} \right].
\]

Inserting the bound on the probability from Lemma 6.45 on the domain of integration we get

\[
\mathbb{P} \left( M(t) - \sqrt{2}t \geq z + y + \sqrt{2}t \right) \leq \rho' \left( \frac{3}{2\sqrt{2}} \ln t + z + y \right) e^{-\sqrt{2}(\frac{3}{2\sqrt{2}} \ln t + z + y)} e^{-y^2/2t}. \tag{7.3.23}
\]

Inserting this into (7.3.22), this is bounded by

\[
\rho'' t^{-3/2} \int_0^{A_1 \sqrt{t}} \left( \frac{3}{2\sqrt{2}} \ln t + z + y \right) e^{-y^2/2t} dy \leq \rho''' A_1^3, \tag{7.3.24}
\]

Similarly, we have that

\[
\mathbb{P} \left( \exists_{i,k} : \eta_i + x_k^{(i)}(t) - \sqrt{2}t \geq z, \text{ but } \eta_i \leq -A_2 \sqrt{t} \right) \leq \int_{A_2 \sqrt{t}}^\infty \mathbb{P} \left( M(t) - \sqrt{2}t \geq z + y + \sqrt{2}t \right) Cy e^{\sqrt{2}y} dy
\]

\[
\leq \bar{\rho} r^{-3/2} \int_{A_2 \sqrt{t}}^\infty y^2 e^{-y^2/2t} dy = \bar{\rho} \int_{A_2 \sqrt{t}}^\infty y^2 e^{-y^2/2t} dy,
\]

which manifestly tends to zero as \( A_2 \uparrow \infty \). \( \square \)

We see that there is a close analogy to Lemma 7.4. How should we interpret the auxiliary process? Think of a very large time \( t \), and go back to time \( t - r \). A that time, there is a certain distribution of particles which all lie well below \( \sqrt{2}(t - r) \) by an order of \( \sqrt{r} \). From those a small fraction will have offspring that reach the
excessive size $\sqrt{2}r + \sqrt{r}$ and thus contribute to the extremes. The selected particles form the Poisson process, while their offspring, which are now BBMs conditioned to be extra large, form the clusters.

### 7.3.4 The extremal process seen from the cluster extremes

We finally come to yet another description of the extremal process. Here we start with the Poisson process of Proposition 7.13 and look at the law of the clusters "that made it up to there". Now we know that the point in the auxiliary process all came from the $\sqrt{t}$ window around $-\sqrt{2t}$. Thus we may expect the clusters to look like BBM that was bigger than $\sqrt{2t}$. Therefore we define the process

$$E_t = \sum_{k \leq n(t)} \delta_{x_k(t) - \sqrt{2t}}. \quad (7.3.26)$$

Obviously, the limit of such a process must be trivial, since the probability that the maximum of BBM shifted by $-\sqrt{2t}$ does not drift to $-\infty$ is vanishing. However, conditionally on the event \(\{\max x_k(t) - \sqrt{2t} \geq 0\}\), the process $E_t$ does converge to a well defined point process $E = \sum_j \delta_{\xi_j}$ as $t \uparrow \infty$.

We may then define the point process of the gaps

$$D_t \equiv \sum_k \delta_{x_k(t) - \max_{j \leq n(t)} x_j(t)}, \quad (7.3.27)$$

and

$$D = \sum_j \delta_{\Delta_j}, \quad \Delta_j \equiv \xi_j - \max_i \xi_i, \quad (7.3.28)$$

where $\xi_i$ are the atoms of the limiting process $E$. Note that $D$ is a point process on $(\infty, 0]$ with an atom at 0.

**Theorem 7.15.** Let $\mathcal{D}_Z$ be as in (7.3.15) and let \(\{\mathcal{D}^{(i)}, i \in \mathbb{N}\}\) be a family of independent copies of the gap-process (7.3.28). Then the point process $E_t = \sum_{k \leq n(t)} \delta_{x_k(t) - m(t)}$ converges in law as $t \uparrow \infty$ to a Poisson cluster point process $E$ given by

$$E \equiv \lim_{t \uparrow \infty} E_t \overset{\text{law}}{=} \sum_{i,j} \delta_{p_i + \Delta_j^{(i)}}. \quad (7.3.29)$$

This theorem looks quite reasonable. The only part that may be surprising is that all the $\Delta_j$ have the same law, since a priori we only know that they come from a $\sqrt{t}$ neighbourhood below $\sqrt{2t}$. This is the content of the next theorem.

**Theorem 7.16.** Let $x \equiv -a \sqrt{t} + b$ for some $a > 0, b \in \mathbb{R}$. The point process

$$\sum_{k \leq n(t)} \delta_{x_k(t) - \sqrt{2t}} \quad (7.3.30)$$
converges in law under \( P \left( \cdot \mid \{ x + \max_k x_k(t) - \sqrt{2}t > 0 \} \right) \), as \( t \uparrow \infty \) to a well-defined point process \( \overline{E} \). The limit does not depend on \( a \) and \( b \), and the maximum of \( \overline{E} \) shifted by \( x \) has the law of an exponential random variable of parameter \( \sqrt{2} \).

**Proof.** Set \( \max \overline{E}_t \equiv \max_i \{ x_i(t) - \sqrt{2}t \} \). We first show that, for \( z \geq 0 \),

\[
\lim_{t \uparrow \infty} P \left( x + \max \overline{E}_t > z \left| x + \max \overline{E}_t > 0 \right. \right) = e^{-\sqrt{2}z} \tag{7.3.31}
\]

for \( X > 0 \). This follows from the fact that the conditional probability is just

\[
P \left( x + \max \overline{E}_t > X \right) \over P \left( x + \max \overline{E}_t > 0 \right), \tag{7.3.32}
\]

and the numerator and denominator can be well approximated by the functions \( \psi(r, t - x + \sqrt{2}t) \) and \( \psi(r, -x + \sqrt{2}t) \), respectively. Using Lemma 7.3 for these, we get the assertion (7.3.31).

Second, we show that for any function \( \phi \) that is continuous has support bounded from below, the limit of

\[
E \left[ \exp \left( - \int (x + z) \overline{E}_t(dz) \right) \left| x + \max \overline{E}_t > 0 \right. \right] \tag{7.3.33}
\]

exists and is independent of \( x \). The proof is just a tick more complicated than before, but relies again on the properties of the functions \( \psi \). I will skip the details. \( \square \)

We now can proof Theorem 7.15.

**Proof (Proof of Theorem 7.15).** The Laplace functional of the process in (9.3.1) of the extremal process clearly is given by

\[
E \left[ \exp \left( -CZ \int_{-\infty}^{\infty} E \left[ 1 - e^{-\frac{1}{2} \phi(y + z) \rho(dz)} \sqrt{2}e^{\sqrt{2}y} dy \right] \right) \right] \tag{7.3.34}
\]

for the point process \( \mathcal{D} \) defined in (7.3.28). We show that the limiting Laplace functional of the auxiliary process can be written in the same form. Then Theorem 7.15 follows from Theorem ???. The Laplace functional of the the auxiliary process, is given by

\[
\lim_{r \uparrow \infty} \Psi_t(x) = \lim_{r \uparrow \infty} E \left[ \exp \left( - \sum_{l,k} \phi(\eta_l + \frac{1}{\sqrt{2}} \ln Z + x^{(j)}_{l,k}(t) - \sqrt{2} t) \right) \right]. \tag{7.3.35}
\]

Using the form for the Laplace transform of a Poisson process we have for the right side
\[
\lim_{t \to \infty} \mathbb{E} \left[ \exp \left( - \sum \phi(\eta_i + \frac{1}{2^k} \ln Z + x_k^{(i)}(t) - \sqrt{2t}) \right) \right] = E \left[ \exp \left( -Z \lim_{t \to \infty} \int_{-\infty}^{0} E \left[ \frac{\phi(x+y)\mathcal{B}_t(dx)}{-\int \phi(x+y)\mathcal{B}_t(dx)} \right] \sqrt{\frac{2}{\pi}}(-y)e^{-\sqrt{2y}dy} \right] .
\]

Define
\[
\mathcal{D}_t = \sum_{t \leq m(t)} \delta_{x_j(t) - \max_{y \in \Omega(t)} x_j(t)} .
\]

The integral on the right-hand side equals
\[
\lim_{t \to \infty} \int_{-\infty}^{0} E \left[ f \left( \int \phi(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \right) \right] \sqrt{\frac{2}{\pi}}(-y)e^{-\sqrt{2y}dy} \]
with \( f(x) = 1 - e^{-x} \). By Proposition 7.14, there exist \( A_1 \) and \( A_2 \) such that
\[
\int_{-\infty}^{0} E \left[ f \left( \int \phi(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \right) \right] \sqrt{\frac{2}{\pi}}(-y)e^{-\sqrt{2y}dy} \]
\[
= \int_{-A_1\sqrt{t}}^{-A_2\sqrt{t}} E \left[ f \left( \int \phi(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \right) \right] \sqrt{\frac{2}{\pi}}(-y)e^{-\sqrt{2y}dy},
\]
\[
+ \Omega_t(A_1,A_2)
\]
where the error term satisfies \( \lim_{t \to \infty} \sup_{t \geq 0} \Omega_t(A_1,A_2) = 0 \). Let \( m_\phi \) be the minimum of the support of \( \phi \). Note that
\[
f \left( \int \phi(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \right)
\]
is zero when \( y + \max \mathcal{B}_t < m_\phi \), and that the event \( \{ y + \max \mathcal{B}_t = m_\phi \} \) has probability zero. Therefore,
\[
\mathbb{E} \left[ f \left( \int \phi(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \right) \right]
\]
\[
= \mathbb{E} \left[ f \left( \int \phi(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \right) \mathbb{1}_{\{y + \max \mathcal{B}_t > m_\phi\}} \right]
\]
\[
= \mathbb{E} \left[ f \left( \int \phi(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \right) \left| y + \max \mathcal{B}_t > m_\phi \right\} \mathbb{P} \left( y + \max \mathcal{B}_t > m_\phi \right) \right].
\]

One can show (see Corollary 4.12 in [6]) that the conditional law of the pair \( (\mathcal{D}_t, y + \max \mathcal{B}_t) \) given \( \{ y + \max \mathcal{B}_t > m_\phi \} \) converges, as \( t \to \infty \), to a pair of independent random variables, where the limit of \( a + \max \mathcal{B}_t \) is exponential with parameter \( \sqrt{2} \) (see (7.3.31)). Moreover, the convergence is uniform in \( y \in [-A_1\sqrt{t} - A_2\sqrt{t}] \). This implies the convergence of the random variable \( f(z+y + \max \mathcal{B}_t) \mathcal{D}_t(dz) \).
Therefore, writing the expectation over the exponential r.v. explicitly, we obtain
\[
\lim_{t \to \infty} \mathbb{E} \left[ f \left( \int \phi (z + s + \max \mathcal{E}_t) \mathcal{D}_t \, dz \right) \right] e^{\max \mathcal{E}_t} = e^{\sqrt{2} m \phi} \int_{m \phi}^{\infty} \mathbb{E} \left[ f \left( \int \phi (z + s) \mathcal{D}_t \, dz \right) \right] \sqrt{2} e^{-\sqrt{2} s} \, ds. 
\] (7.3.43)

On the other hand,
\[
\int_{-A_2 \sqrt{t}}^{-A_1 \sqrt{t}} \mathbb{P} \left( y + \max \mathcal{E}_t > m \phi \right) \sqrt{\frac{2}{\pi}} (-y) e^{-\sqrt{2} y} \, dy = C e^{-\sqrt{2} m \phi} + \Omega_t (A_1, A_2) 
\] (7.3.44)

by Lemma 7.6, where \( \lim_{A_1 \to 0, A_2 \to \infty} \limsup_{t \to \infty} \Omega_t (A_1, A_2) = 0 \), using the same approximation as in (7.3.39).

Combining (7.3.44), (7.3.43) and (7.3.42), one sees that (7.3.36) converges to
\[
\mathbb{E} \left[ \exp \left( -CZ \int_{-\infty}^{r} \mathbb{E} \left[ 1 - e^{-\int \phi (y + z) \mathcal{D}_t \, dz} \right] \sqrt{2} e^{-\sqrt{2} y} \, dy \right) \right], 
\] (7.3.45)

which is by (7.3.35) the limiting Laplace transform of the extremal process of branching Brownian motion: this shows (7.3.34) and concludes the proof of Theorem 7.15. \( \square \)

The properties of BBM conditioned to exceed their natural threshold were already described in detail by Chauvin and Rauoult [28]. There will be one branch (the spine) that exceeds the level \( \sqrt{2} t \) by an exponential random variable of parameter \( \sqrt{2} \) (see the preceding proposition). The spine is very close to a straight line of slope \( \sqrt{2} \).

From this spine, ordinary BBM’s branch off at Poissonian times. Clearly, all the branches that split off at times later than \( r \) before the end-time, will reach at most the level
\[
\sqrt{2} (t - r) + \sqrt{2} r - \frac{3}{2\sqrt{2}} \ln r = \sqrt{2} t - \frac{3}{2\sqrt{2}} \ln r. 
\] (7.3.46)

Seen from the top, i.e. \( \sqrt{2} t \), this tends to \( -\infty \) as \( r \to \infty \). Thus only branches that are created "a finite time” before the end-time \( r \) remain visible in the extremal process. This does, of course, correspond perfectly to the observation in Theorem 7.11 that implies that all the points visible in the external process have a common ancestor at a finite time before the end. As a matter of fact, from the observation above one can prove Theorem 7.11 easily.
In this chapter we present an extension of the convergence of the extremal process of BBM. This material is taken from Bovier and Hartung [16].

8.1 Introduction

We have seen in Chapter 2 that in extreme value theory (see e.g. [57]) it is customary to give description of extremal processes that also encode the locations of the extreme points ("complete Poisson convergence"). We want to obtain an analogous result for BBM. Now our "space" is the Galton-Watson tree, or more precisely its "leaves" at infinity. This requires a construction of this object. In the case of deterministic binary branching at integer times, the leaves of the tree at time $n$ are naturally labelled by sequences $\sigma^n \equiv (\sigma_1, \sigma_2, \ldots, \sigma_n)$, with $\sigma_\ell \in \{0, 1\}$. These sequences can be naturally mapped into $[0, 1]$ via

$$\sigma^n \mapsto \sum_{\ell=1}^n \sigma_\ell 2^{-\ell-1} \in [0, 1].$$

(8.1.1)

Moreover, the limit, as $n \uparrow \infty$ of the image of this map is $[0, 1]$. We now want to do the same for a general Galton-Watson tree.

8.1.1 The embedding

Our goal is to define a map $\gamma: \{1, \ldots, n(t)\} \to \mathbb{R}_+$ in such a way that it encodes the genealogical structure of the underlying supercritical Galton-Watson process. Recall the we write $\hat{I}^k(t)$ for the multi-index that labels the $k$-th particle at time $t$ and $x_k(t)$ for the position of that particle. *multi-indices*. To this end, recall the labelling of the Galton-Watson tree through multi-indices presented in Section 4.5. Then define
\[ \gamma(i) \equiv \gamma(i(t)) \equiv \sum_{j=1}^{W(t)} i_j(t)e^{-t_j}. \quad (8.1.2) \]

For a given \( i \), the function \( \gamma(i(t)), t \in \mathbb{R}_+ \) describes a trajectory of a particle in \( \mathbb{R}_+ \). Thus to any "particle" at time \( t \) we can now associate the position on \( \mathbb{R} \times \mathbb{R}_+ \), \((x_i(t), \gamma(i(t)))\). The important point is that for any given particle, this trajectory converges to some point \( \gamma(i) \in \mathbb{R}_+ \), as \( t \uparrow \infty \), almost surely. Hence also the sets \( \gamma(\tau(t)) \) converge, for any realisation of the tree, to some (random) set \( \gamma(\tau(\infty)) \). It is easy to see that \( \gamma(\tau(\infty)) \) is a Cantor set.

### 8.1.2 The extended convergence result

Using the embedding \( \gamma \) defined in the previous section, we now state the following theorem. We use the notation from the previous chapter.

**Theorem 8.1.** The point process \( \tilde{E}_t \equiv \sum_{k=1}^{n(t)} \delta_{\gamma(x_i(t), x_i(t) - m(t))} \to \tilde{E} \) on \( \mathbb{R}_+ \times \mathbb{R} \), as \( t \uparrow \infty \), where

\[ \tilde{E} = \sum_{i,j} \delta_{(q_i,p_i) + (0, \Delta_j^{(i)})}, \quad (8.1.3) \]

where \((q_i,p_i)_{i \in \mathbb{N}}\) are the atoms of a Cox process on \( \mathbb{R}_+ \times \mathbb{R} \) with intensity measure \( Z(dv) \times C e^{-\sqrt{2}x}dx \), where \( Z(dv) \) is a random measure on \( \mathbb{R}_+ \), characterised in Lemma 8.4, and \( \Delta_j^{(i)} \) are the atoms of the independent and identically distributed point processes \( \Delta_j \) from Theorem 7.15.

**Remark 8.2.** The nice feature of the process \( \tilde{E}_t \) is that it allows to visualise the different clusters \( \Delta_j \) corresponding to the different point of the Poisson process of cluster extremes. In the process \( \sum_{k=1}^{n(t)} \delta_{x_i(t) - m(t)} \) considered in earlier work, all these points get superimposed and cannot be disentangled. In other words, the process \( \tilde{E} \) encodes both the values and the (rough) genealogical structure of the extremes of BBM.

**Remark 8.3.** A very similar result is proven by Biskup and Louidor for the discrete Gaussian free field \([13, 14]\). Their result provided the motivation to prove the corresponding theorem for BBM.

The measure \( Z(dv) \) is an interesting object in itself. For \( v, r \in \mathbb{R}_+ \) and \( t > r \), we define

\[ Z(v, r, t) = \sum_{j \leq n(t)} (\sqrt{2}t - x_j(t))e^{\sqrt{2}(x_j(t) - \sqrt{2}t)}I_{\gamma(i(t)) \leq v}, \quad (8.1.4) \]

which is a truncated version of the usual derivative martingale \( Z_t \). In particular, observe that \( Z(\infty, r, t) = Z_t \).
Lemma 8.4. For each \( v \in \mathbb{R}_+ \) the limit \( \lim_{r \to \infty} \lim_{t \to \infty} Z(v, r, t) \) exists almost surely. Set
\[
Z(v) \equiv \lim_{r \to \infty} \lim_{t \to \infty} Z(v, r, t).
\]
Then \( 0 \leq Z(v) \leq Z \), where \( Z \) is the limit of the derivative martingale. Moreover, almost surely, \( Z(v) \) is monotone increasing in \( u \) and continuous.

The measure \( Z(v) \) is the analogue of the corresponding "derivative martingale measure" studied in Duplantier et al [33, 34] and Biskup and Louidor [13, 14] in the context of the Gaussian free field. For a review, see Rhodes and Vargas [66]. The objects are examples of what is known as multiplicative chaos that was introduced by Kahane [48].

8.2 Properties of the embedding

We need the three basic properties of \( \gamma \). The first lemma states that the map \( \gamma(R^k(t)) \) is well approximated by \( \gamma(R^k(r)) \), when \( r \) is large, but small compared to \( t \), provided \( t \) is large enough. More precisely:

**Lemma 8.5.** Let \( D = [\overline{D}, \overline{D}] \subset \mathbb{R} \) be a compact interval Define, for \( 0 \leq r < t < \infty \), the events
\[
\omega_{r,t}^D(D) = \left\{ \forall k \text{ with } x_k(t) - m(t) \in D : \gamma(R^k(t)) - \gamma(R^k(r)) \leq e^{-r/2} \right\}.
\]

For any \( \varepsilon > 0 \) there exists \( 0 \leq r(D, \varepsilon) < \infty \) such that, for any \( r > r(D, \varepsilon) \) and \( t > 3r \)
\[
\mathbb{P}(\omega_{r,t}^D(D)) < \varepsilon.
\]

**Proof.** Let \( \varepsilon > 0 \). Then, by Theorem 2.3 of [33] there exists for each \( \varepsilon > 0 \) \( r_1 < \infty \) such that, for all \( t > 3r_1 \)
\[
\mathbb{P}(\omega_{r,t}^D(D)) \leq \mathbb{P}(\exists k : x_k(t) - m(t) \in D, \forall s \in [r_1, t - r_1] : x_k(s) \leq \overline{D} + E_{i,\alpha}(s) \text{ but } \gamma(R^k(t)) - \gamma(R^k(r)) > e^{-r/2} + \varepsilon/2, \quad (8.2.3)
\]

where \( 0 < \alpha < \frac{1}{2} \) and \( E_{i,\alpha}(s) = \gamma(m(t) - f_{i,\alpha}(s)) \) and \( f_{i,\alpha} = (s \wedge (t - s))^\alpha \). Using the "many-to-one lemma" (see Theorem 8.5 of [42]), the probability in (8.2.3) is bounded from above by
\[
e^{\varepsilon} \mathbb{P}
left( x(t) \in m(t) + D, \forall s \in [r_1, t - r_1] : x(s) \leq \overline{D} + E_{i,\alpha}(s) \text{ but } \sum_j m_j e^{-j} \mathbb{1}_{i_j \in [r_1]} > e^{-r/2} \right),
\]

where \( x \) is a standard Brownian motion and \( (\tilde{l}_j, j \in \mathbb{N}) \) are the points of a size-biased Poisson point process with intensity measure \( 2dx \) independent of \( x \), \( m_j \) are independent random variables, uniformly distributed on \( \{0, \ldots, \tilde{l}_j - 1\} \), where finally \( \tilde{l}_j \) are i.i.d. according to the size-biased offspring distribution, \( \mathbb{P}(\tilde{l}_j = k) = \frac{k \nu}{2} \). Due to
inddependence, and since \( m_j \leq \hat{I}_j \), the expression (8.2.4) is bounded from above by

\[
e^{-\epsilon t} \mathbb{P}\left( x(t) \in m(t) + D, \forall s \in [r_1, t - r_1]: x(s) \leq D + E_1, \alpha (s) \right) \times \mathbb{P}\left( \sum_j (\hat{I}_j - 1) e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} > e^{-t/2} \right).
\]

(8.2.5)

The first probability in (8.2.5) is bounded by

\[
\mathbb{P}\left( x(t) \in m(t) + D, \forall s \in [r_1, t - r_1]: x(s) - \frac{s}{t} x(t) \leq D - D - f_1, \alpha (s) \right).
\]

(8.2.6)

Using that \( \xi (s) \equiv x(s) - \frac{s}{t} x(t) \) is a Brownian bridge from 0 to 0 in time \( t \) that is independent of \( x(t) \), (8.2.6) equals

\[
\mathbb{P}\left( x(t) \in m(t) + D \right) \mathbb{P}\left( \forall s \in [r_1, t - r_1]: \xi (s) \leq D - D - f_1, \alpha (s) \right)
\]

\[
\leq \mathbb{P}\left( x(t) \in m(t) + D \right) \mathbb{P}\left( \forall s \in [r_1, t - r_1]: \xi (s) \leq D - D \right).
\]

(8.2.7)

Using Lemma 3.4 of [3] to bound the last factor of (8.2.7) we obtain that (8.2.7) is bounded from above by

\[
\kappa \frac{r_1}{t - 2r_1} \mathbb{P}\left( x(t) \in m(t) + D \right) \mathbb{P}\left( \sum_j (\hat{I}_j - 1) e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} > e^{-t/2} \right).
\]

(8.2.8)

where \( \kappa < \infty \) is a positive constant. Using this as an upper bound for the first probability in (8.2.5) we can bound (8.2.5) from above by

\[
e^{-\epsilon t} \kappa \frac{r_1}{t - 2r_1} \mathbb{P}\left( x(t) \in m(t) + D \right) \mathbb{P}\left( \sum_j (\hat{I}_j - 1) e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} > e^{-t/2} \right).
\]

(8.2.9)

By (5.25) of [3] (resp. an easy Gaussian computation) this is bounded from above by

\[
C \kappa \frac{r_1}{t - 2r_1} \mathbb{P}\left( \sum_j (\hat{I}_j - 1) e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} > e^{-t/2} \right).
\]

(8.2.10)

for some positive constant \( C < \infty \). Using the Markov inequality, (8.2.10) is bounded from above by

\[
C \kappa \frac{r_1}{t - 2r_1} e^{t/2} \mathbb{E}\left( \sum_j (\hat{I}_j - 1) e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} \right).
\]

(8.2.11)

We condition on the \( \sigma \)-algebra \( F \) generated by the Poisson points. Using that \( \hat{I}_j \) is independent of the Poisson point process \( (\hat{I}_j)_j \) and \( \sum_j e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} \) is measurable with respect to \( F \) we obtain that (8.2.11) is equal to

\[
C \kappa \frac{r_1}{t - 2r_1} e^{t/2} \mathbb{E}\left( \mathbb{E}\left( \sum_j (\hat{I}_j - 1) e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} \big| F \right) \right)
\]

\[
= C \kappa \frac{r_1}{t - 2r_1} e^{t/2} \mathbb{E}\left( \sum_j e^{-\epsilon j} \mathbb{1}_{\hat{I}_j \in [r_1]} \mathbb{E}\left( (\hat{I}_j - 1) \big| F \right) \right).
\]

(8.2.12)

Since \( \mathbb{E}(I_j - 1) = \sum_k \frac{1}{2}(k - 1)k = K/2 < \infty \) we have that (8.2.12) is equal to
\[ C \kappa K / 2 \frac{t r_1}{t - 2 r_1} e^{r/2} E \left( \sum_j e^{-t} \mathbb{1}_{t_j \in [r]} \right). \quad (8.2.13) \]

By Campbell’s theorem (see e.g. [52]), (8.2.13) is equal to
\[ C \kappa K / 2 \frac{t r_1}{t - 2 r_1} e^{r/2} \int_r^t e^{-x} 2 dx \leq C \kappa K / 2 \frac{t r_1}{t - 2 r_1} e^{-r/2}, \quad (8.2.14) \]
which can be made smaller than \( \varepsilon / 2 \) for all \( r \) sufficiently large and \( t > 3r \).

The second lemma ensures that \( \gamma \) maps particles are with a low probability to a very small neighbourhood of a fixed \( a \in \mathbb{R} \).

**Lemma 8.6.** Let \( a \in \mathbb{R}_+ \) and \( D \) as in Lemma ???. Define the event
\[ B_{\gamma}^r(D,a,\delta) = \{ \forall k \text{ with } x_k(t) - m(t) \in D: \gamma(k(t)) \notin [a-\delta,a] \}. \quad (8.2.15) \]
For any \( \varepsilon > 0 \) there exists \( \delta > 0 \) and \( r(a,D,\delta,\varepsilon) \) such that for any \( r > r(a,D,\delta,\varepsilon) \) and \( t > 3r \)
\[ P \left( \left( B_{\gamma}^r(D,a,\delta) \right)^c \right) < \varepsilon. \quad (8.2.16) \]

**Proof.** Following the proof of Lemma 8.5 step by step we arrive at the bound
\[ P \left( \left( B_{\gamma}^r(D,a,\delta) \right)^c \right) \leq C \kappa K / 2 \frac{t r_1}{t - 2 r_1} \sum_j m_j e^{-\frac{t}{2}} \mathbb{1}_{j \in [0,r]} \in [a-\delta,a]. \quad (8.2.17) \]
We rewrite the probability in (8.2.17) in the form
\[ \sum_{i^* = 1}^{\infty} P \left( i^* = \inf \{ i : m_i \neq 0 \}, \sum_{j \geq i^*} m_j e^{-\frac{t}{2}} \mathbb{1}_{j \in [0,r]} \in [a-\delta,a] \right). \quad (8.2.18) \]
Consider first \( P \left( i^* = \inf \{ i : m_i \neq 0 \} \right) \). This probability is equal to
\[ P \left( \forall i \leq i^* : m_i = 0 \text{ and } m_{i^*} \neq 0 \right) = E \left( \left( 1 - \frac{1}{l_{i^*}} \right) \prod_{j=1}^{i^*-1} \frac{1}{l_j} \right). \quad (8.2.19) \]
Using that the \( l_j \) are iid together with the simple bound \( E \left( l_{i^*}^{-1} \right) \leq \frac{1+p_1}{2} \), we see that (8.2.19) is bounded from above by
\[ \left( \frac{1+p_1}{2} \right)^{i^*-1}. \quad (8.2.20) \]
Since \( \frac{1+p_1}{2} < 1 \) by assumption on \( p_1 \) we can choose for each \( \varepsilon' > 0 \ K(\varepsilon') < \infty \) such that
\[ \sum_{i^* = K(\varepsilon')+1}^{\infty} \left( \frac{1+p_1}{2} \right)^{i^*-1} < \varepsilon'. \quad (8.2.21) \]
Hence we bound (8.2.18) by
\[
K(\varepsilon') \leq \sum_{i=1}^{\infty} \mathbb{P} \left( t^* = \inf \{ i : m_i \neq 0 \}, \sum_{j \geq r} m_j e^{-t_j} \mathbb{1}_{t_j \in [0, r]} \right) + \epsilon'. \tag{8.2.22}
\]

We rewrite
\[
\sum_{j \geq r} m_j e^{-t_j} \mathbb{1}_{t_j \in [0, r]} = m_r e^{-t_r} \mathbb{1}_{t_r \in [0, r]} \left( 1 + m_r^{-1} \sum_{j > r} m_j e^{-(t_j-t_r)} \mathbb{1}_{t_j-t_r \in [0, r-r]} \right) \tag{8.2.23}
\]

Next, we estimate the probability that \( \tilde{t}_r \) is large. Observe that \( \tilde{t}_r = \sum_{i=1}^{r} s_i \) where \( s_i \) are iid exponentially distributed random variables with parameter 2. This implies that \( \tilde{t}_r \) is Erlang(2, \( i^* \)). Thus
\[
\mathbb{P}(\tilde{t}_r > r^*) = e^{-2r^*} \sum_{j=0}^{r^*} \frac{(2r^*)^j}{j!}\leq C(K(\varepsilon'))b(2r^*)K(\varepsilon')e^{-2r^*}, \text{ for all } r^* \leq K(\varepsilon). \tag{8.2.24}
\]

Next we want to replace \( \tilde{t}_r \) in the indicator function in (8.2.23) by a non-random quantity \( r^* \), for some \( 0 < \alpha < 1 \), in order to have a bound that depends only on the differences \( t_j - \tilde{t}_r \). Note first that
\[
\sum_{j > r} m_j e^{-(t_j-\tilde{t}_r)} \mathbb{1}_{t_j-\tilde{t}_r \in [0, r-r]} = \sum_{j > r} m_j e^{-(t_j-\tilde{t}_r)} \mathbb{1}_{t_j-\tilde{t}_r \in [r-r^*, r-r]} \leq \sum_{j > r} m_j e^{-(t_j-\tilde{t}_r)} \mathbb{1}_{t_j-\tilde{t}_r \in [r-r^*, r]}.
\tag{8.2.25}
\]

Using the fact that \( m_j \leq \tilde{t}_j - 1 \) for all \( j \) and the Markov inequality, we get that
\[
\mathbb{P} \left( \sum_{j > r} m_j e^{-(t_j-\tilde{t}_r)} \mathbb{1}_{t_j-\tilde{t}_r \in [r-r^*, r]} > e^{-r/2} \right) \leq e^{r/2} \mathbb{E} \left( \sum_{j > r} (\tilde{t}_j - 1) e^{-(t_j-\tilde{t}_r)} \mathbb{1}_{t_j-\tilde{t}_r \in [r-r^*, r]} \right). \tag{8.2.26}
\]

Using Campbell’s theorem as in (8.2.12), we see that the second line in (8.2.26) is equal to
\[
e^{r/2} K/2 \int_{r-r^*}^r e^{-x} dx = K \left( e^{-r/2} e^{-r^*} - e^{-r/2} \right).
\tag{8.2.27}
\]

For any \( \varepsilon' > 0 \), there exists \( r_0 < \infty \), such that for all \( r > r_0 \), the probabilities in (8.2.24) and (8.2.26) are smaller than \( \varepsilon' \). On the the event
\[
\mathcal{D} = \{ t_r \leq r^* \} \cap \left\{ \sum_{j > r} m_j e^{-(t_j-\tilde{t}_r)} \mathbb{1}_{t_j-\tilde{t}_r \in [r-r^*, r]} \leq e^{-r/2} \right\}, \tag{8.2.28}
\]

which has probability at least \( 1 - 2\varepsilon' \), we can bound (8.2.22) in a nice way. Namely, since \( m_r \geq 1 \) by definition and \( m_j \) are chosen uniformly from \( \{0, \ldots, \tilde{t}_j - 1\} \) and independent of \( \{t_j\}_{j \geq 1} \). Moreover, \( \sum_{j > r} m_j e^{-(t_j-\tilde{t}_r)} \mathbb{1}_{t_j-\tilde{t}_r \in [0, r-r^*]} \geq 0 \) is also independent of \( t_{r_0} \). It follows that (8.2.22) is bounded from above by
Inserting this bound into (8.2.31), we get that, for some constant $C > 0$ and $\delta > 0$ small enough, we have that 

\[
|\tilde{t}_r - t_r| \leq C(K(e'))(\delta + e^{-r/2}),
\]

for some constant $C(K(e')) < \infty$. Now the right-hand side of (8.2.34) can be made smaller than $\epsilon'$ by choosing $r$ large enough and $\delta$ small enough.

Collecting the bounds in (8.2.24), (8.2.26) and (8.2.34) implies (8.2.16) if $\epsilon' = \epsilon/4$.

The following lemma asserts that any two points that get close to the maximum of BBM with have distinct images under the map $\gamma$, unless their genealogical distance is large.

**Lemma 8.7.** Let $D \subset \mathbb{R}$ be as in Lemma 8.6. For any $\epsilon > 0$ there exists $\delta > 0$ and $r(\delta, \epsilon)$ such that for any $r > r(\delta, \epsilon)$ and $t > 3r$ 

\[
P \left( \exists i, j \leq n(t) : d(x_i(t), x_j(t)) \leq \epsilon : x_i(t), x_j(t) \in m(t) + D, |\gamma(\tilde{x}(t)) - \gamma(\tilde{y}(t))| \leq \delta \right) < \epsilon.
\]

(8.2.35)
Proof. To control (8.2.35), we first use that, by Theorem 2.1 in [3], for any \( \varepsilon' \), there is \( r_1 < \infty \), such that, for all \( t \geq 3r_1 \), and \( r \leq t/3 \), the event
\[
\{ \exists_{i,j \in \mathcal{N}(t)} \mid d(x_i(t), x_j(t)) \leq (r_1, r), x_i(t), x_j(t) \in m(t) + D \}
\]  
has probability smaller than \( \varepsilon' \). Therefore,
\[
P \left( \exists_{i,j \in \mathcal{N}(t)} \mid d(x_i(t), x_j(t)) \leq r, x_i(t), x_j(t) \in m(t) + D, |\gamma_i(t') - \gamma_i(t')| \leq \delta \right) \leq P \left( \exists_{i,j \in \mathcal{N}(t)} \mid d(x_i(t), x_j(t)) \leq r_1 : x_i(t), x_j(t) \in m(t) + D, |\gamma_i(t') - \gamma_i(t')| \leq \delta \right) + \varepsilon'.
\]  
The nice feature of the probability in the last line is that \( r_1 \) is now independent of \( r \). At the expense of one more \( \varepsilon' \), we can introduce in addition the condition that the paths on \( x_i(t), x_j(t) \) are localised in \( E_{r_1} \) over the interval \([r_2, t - r_2] \), for some \( r_1 < r_2 < \infty \), independent of \( t \). Then a second moment estimate (also known as the many-to-two lemma), shows that
\[
P \left( \exists_{i,j \in \mathcal{N}(t)} \mid d(x_i(t), x_j(t)) \leq r_1 : x_i(t), x_j(t) \in m(t) + D, |\gamma_i(t') - \gamma_i(t')| \leq \delta \right) \leq e^{2n} K P \left( \exists_{i \in \mathcal{N}(t)} \mid d(x_i(t), x_i(t)) \leq r_1 : x_i(t), x_i(t) \in m(t) + D, \forall s \in [r_1, t - r_2], \right.
\]
\[
\exists_i \left( t_1^{(i)}(s), t_2^{(i)}(s) \right) \leq D + E_{r_1} \kappa, k = 1, 2, |\gamma_i(t^{(i)} - \gamma_i(t^{(i)})) \leq \delta \right) + \varepsilon'.
\]  
where we write \( x_i^{(k)}(t) = x_i(r_1) + x^{(k)}(t - r_1) \), etc., and \( D \) is a finite enlargement of \( D \) such that \( D + x_i(r_1) \subset D \) with probability at least \( 1 - \varepsilon' \), and \( D \) is the supremum of \( D \). Using independence of the branches \( x^{(k)} \) and the same arguments as in Lemma 8.5, we see that the probability in the last line is bounded from above by
\[
\left( C \sqrt{\frac{t_2}{t - r_2}} \right)^2 P \left( \left| \gamma_i(t^{(1)}(r_1)) - \gamma_i(t^{(2)}(r_1)) + \sum_k m_k e^{t^{(k)} - t} \right| \leq \delta \right),
\]  
where \( (\tilde{t}_k, k \in \mathbb{N}) \) and \( (\tilde{t}_k, k' \in \mathbb{N}) \) are the points of independent Poisson point processes with intensity \( 2d \) restricted to \([r_1, t] \). Moreover, \( \tilde{t}_k, \tilde{t}_k' \) are i.i.d. according to the size-biased offspring distribution and \( m_k \) resp. \( m_k' \) are uniformly distributed on \([0, \ldots, I_k - 1] \) resp. \([0, \ldots, I_k' - 1] \). We rewrite (8.2.39) as
\[
P \left( \sum_k m_k e^{-t} \mathbb{I}_{k \in [r_1, t]} \in \gamma_i(t^{(1)}(r_1)) - \gamma_i(t^{(2)}(r_1)) + \sum_k m_k e^{-t} \mathbb{I}_{k' \in [r_1, t]} + [-\delta, \delta] \right).
\]  
As in (8.2.18) we rewrite the probability in (8.2.40) as
Proposition 8.8. Let $x^{(j)}$ be independent BBM's. Then

\[
\sum_{k \geq t} m_k^j e^{-\gamma_k^j} \in \gamma_t(\mathbf{1}^j(r_1)) - \gamma_t(\mathbf{1}^j(r_1)) + \sum_k m_k^i e^{-\gamma_k^i} + [-\delta, \delta].
\]

Due to the independence of $(\gamma_k^j, k \in \mathbb{N})$ and $(\gamma_k^j, k' \in \mathbb{N})$ we can proceed as with (8.2.18) in the proof of Lemma 8.6 to make (8.2.41) as small as desired by choosing $\delta$ small enough. The prefactor in (8.2.39) tends to a constant as $t \uparrow \infty$, and the additional prefactor from (8.2.38) is independent of $t$ and $\delta$. This implies the assertion of Lemma 8.7.

### 8.3 The $q$-thinning

The proof of the convergence of $\sum_{i=1}^{n(t)} \delta_{\gamma(i(t)), x_{i(t)} - m(t)}$ comes in two main steps. In a first step, we show that the points of the local extrema converge to the desired Poisson point process. Recall from Section 7.3.2 the process $\Theta^i(t)$ defined in (7.3.12).

For $r_d \in \mathbb{R}^+$ and $t > 3r_d$, we can write, with $R_t = m(t) - m(t - r_d) - \sqrt{2}r_d = o(1)$, we have

\[
\Theta^i(t) \overset{D}{=} \sum_{r_d} \delta_{j(t) - \sqrt{2}r_d + M_j(t - r_d) + R_t},
\]

where $M_j(t - r_d) \equiv \max_{k \geq n_j(t - r_d)} x_k^{(j)}(t - r_d) - m(t - r_d)$ and $x^{(j)}$ independent BBM's. Then

**Proposition 8.8.** Let $x_{i_k}(t)$ denote the atoms of the process $\Theta^i(t)$. Then

\[
\lim_{r_d \to 0} \lim_{t \to \infty} \sum_{i_k} \delta_{\gamma(i_k(t)), x_{i_k(t)}(t)} \overset{D}{=} \delta_{\hat{t}},
\]

where $(q_i, p_i) \in \mathbb{N}$ are the points of the Cox process $\hat{t}$ with intensity measure $Z(dv) \times Ce^{-\sqrt{2}v}dv$ with the random measure $Z(dv)$ defined in (8.1.5). Moreover,

\[
\lim_{r_d \to 0} \lim_{t \to \infty} \sum_{i_k} \delta_{\gamma(i(t)) - \sqrt{2}r_d + M_j} \overset{D}{=} \delta_{\hat{t}},
\]

where $M_j$ are i.i.d. and distributed with the limit law of the maximum on BBM.

The proof of Proposition 8.8 relies in Lemma 8.4 which we now prove.

**Proof (Proof of Lemma 8.4).** For $v, r \in \mathbb{R}^+$ fixed, the process $Z(v, r, t)$ defined in (8.1.4) is a martingale in $t > r$ (since $Z(\infty, r, t)$ is the derivative martingale and $\mathbf{1}_{\delta_i(\nu(r)) \leq v}$ does not depend on $t$). To see that $Z(v, r, t)$ converges a.s. as $t \uparrow \infty$, note that
Lemma 8.7 there exists increasing in the almost sure limit of the derivative martingale. Thus \( \lim_{t \to \infty} \)

Here \( Z_i(t), i \in \mathbb{N} \) are iid copies of the derivative martingale, and \( Y_i(t) \) are iid copies of the McKean martingale. From Section 5.6 we know that \( Y_i(t) \to 0 \), a.s., while \( Z_i(t) \to Z(t) \), a.s., where \( Z(t) \) are non-negative r.v.'s. Hence

\[
\lim_{t \to \infty} Z(v, r, t) = Z(v, r) = \sum_{i=1}^{n(r)} e^{\sqrt{2}(x_i(t) - \sqrt{2}r)} \left( \sum_{j=1}^{n(r)} \left( \sqrt{2}(t - r) - x_j(t) \right) e^{\sqrt{2}(x_j(t) - \sqrt{2}r)} \right) \\
= \sum_{i=1}^{n(r)} \mathbb{1}_{Y_i(t) \leq v} \sum_{i=1}^{n(r)} \mathbb{1}_{Y_i(t) \leq v} e^{\sqrt{2}(x_i(t) - \sqrt{2}r)} \left( \sqrt{2}(t - r) - x_i(t) \right) Y_i(t) \\
+ \sum_{i=1}^{n(r)} \mathbb{1}_{Y_i(t) \leq v} e^{\sqrt{2}(x_i(t) - \sqrt{2}r)} Z_i(t). \tag{8.3.4}
\]

where \( Z(t), i \in \mathbb{N} \) are iid copies of \( Z \). To show that \( Z(v, r) \) converges, as \( r \uparrow \infty \), we go back to (8.1.4). Note that for fixed \( v \), \( \mathbb{1}_{Y_i(t) \leq v} \) is monotone decreasing in \( r \).

On the other hand, we have seen (5.6.8) that \( \min_{v \leq \sqrt{2}r} \left( \sqrt{2}r - x_i(t) \right) \to +\infty \), almost surely, as \( t \uparrow \infty \). Therefore, the part of the sum in (8.1.4) that involves negative terms (namely those for which \( x_i(t) > \sqrt{2}r \)) converges to zero, almost surely. The remaining part of the sum is decreasing in \( r \), and this implies that the limit, as \( t \uparrow \infty \), is monotone decreasing almost surely. Moreover, \( 0 \leq Z(v, r) \leq Z \), a.s., where \( Z \) is the almost sure limit of the derivative martingale. Thus \( \lim_{r \to \infty} Z(v, r) = Z(v) \) exists. Finally, \( 0 \leq Z(v) \leq Z \) and \( Z(v) \) is an increasing function of \( v \) because \( Z(v, r) \) is increasing in \( v \), a.s., for each \( r \).

To show that \( Z(du) \) is nonatomic, fix \( \epsilon, \delta > 0 \) and let \( D \subset \mathbb{R} \) be compact. By Lemma 8.7 there exists \( r_i(\epsilon, \delta) \) such that, for all \( r > r_i(\epsilon, \delta) \) and \( t > 3r \),

\[
\mathbb{P} \left( \exists_{i \leq m(t)} : d(x_i(t), x_j(t)) \leq r, x_i(t), x_j(t) \in m(t) + D, |Y_i(t) - Y_j(t)| \leq \delta \right) < \epsilon. \tag{8.3.6}
\]

Rewriting (8.3.6) in terms of the thinned process \( \mathcal{E}(r/t)(t) \) gives

\[
\mathbb{P} \left( \exists_{i \leq m(t)} : x_i(t) \in m(t) + D, |\mathcal{E}(r/t)(t) - \mathcal{E}(r/t)(t)| \leq \delta \right) \leq \epsilon. \tag{8.3.7}
\]

Assuming for the moment that \( \mathcal{E}(r/t)(t) \) converges as claimed in Proposition 8.8 this implies that for any \( \epsilon > 0 \), for small enough \( \delta > 0 \),

\[
\mathbb{P} \left( \exists \delta > 0 : \exists i \neq j : |r_i - r_j| < \delta \right) < \epsilon. \tag{8.3.8}
\]
This could not be true if $Z(du)$ had an atom. This proves Lemma 8.4 provided we can show convergence of $\sigma^{(t'\ell)}(t)$.

The proof of Proposition 8.8 uses the properties of the map $\gamma$ obtained in Lemma 8.5 and 8.6. In particular, we use that, in the limit as $t \uparrow \infty$, the image of the extremal particles under $\gamma$ converges and that essentially no particle is mapped too close to the boundary of any given compact set. Having these properties at hand we can use the same procedure as in the proof of Proposition 5 in [4]. Finally, we use Lemma 8.4 to deduce Proposition 8.8.

Proof (Proof of Proposition 8.8). We show the convergence of the Laplace functionals. Let $\phi : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$ be a measurable function with compact support. For simplicity we start by looking at simple functions of the form

$$\phi(x,y) = \sum_{i=1}^{N} a_i \mathbf{1}_{A_i \times B_i}(x,y),$$

where $A_i = [\bar{A}_i, A_i]$ and $B_i = [\bar{B}_i, B_i]$ for $N \in \mathbb{N}$, $i = 1, \ldots, N$, $a_i, \bar{A}_i, A_i \in \mathbb{R}_+$, and $\bar{B}_i, B_i \in \mathbb{R}$. The extension to general functions $\phi$ then follows by monotone convergence. For such $\phi$, we consider the Laplace functional

$$\Psi_\ell(\phi) \equiv \mathbb{E} \left[ \exp \left( - \sum_{k=1}^{n_\ell(t)} \phi \left( \gamma_i(t), \tilde{x}_{i_k}(t) \right) \right) \right].$$

The idea is that the function $\gamma$ only depends on the early branchings of the particle. To this end we insert the identity

$$1 = \mathbf{1}_{\mathcal{M}_{\ell}(\text{supp} \phi)} + \mathbf{1}_{(\mathcal{M}_{\ell}(\text{supp} \phi))^c}$$

into (8.3.10), where $\mathcal{M}_{\ell}$ is defined in (8.2.1), and by $\text{supp} \phi$ we mean the support of $\phi$ with respect to the second variable. By Lemma 8.5 we have that, for all $\varepsilon > 0$, there exists $r_\varepsilon$ such that, for all $r > r_\varepsilon$,

$$\mathbb{P} \left( (\mathcal{M}_{\ell}(\text{supp} \phi))^c \right) < \varepsilon,$$  

uniformly in $t > 3r$. Hence it suffices to show the convergence of

$$\mathbb{E} \left[ \exp \left( - \sum_{k=1}^{n_\ell(t)} \phi \left( \gamma_i(t), \tilde{x}_{i_k}(t) \right) \right) \mathbf{1}_{\mathcal{M}_{\ell}(\text{supp} \phi)} \right].$$

We introduce yet another identity into (8.3.13), namely

$$1 = \mathbf{1}_{\bigcap_{i=1}^{N} (\mathcal{M}_{\ell}(\text{supp} \phi, \bar{A}_i)) \cap \mathcal{M}_{\ell}(\text{supp} \phi, \bar{A}_i)} + \mathbf{1}_{(\mathcal{M}_{\ell}(\text{supp} \phi, \bar{A}_i)) \cap (\mathcal{M}_{\ell}(\text{supp} \phi, \bar{A}_i))^c},$$

where we use the shorthand notation $\mathcal{M}_{\ell}(\text{supp} \phi, \bar{A}_i) \equiv \mathcal{M}_{\ell}(\text{supp} \phi, \bar{A}_i, e^{-t/2})$ ( recall (8.2.15)). By Lemma 8.6 there exists for all $\varepsilon > 0 \bar{r}_\varepsilon$ such that for all $r > \bar{r}_\varepsilon$
uniformly in $t > 3r$

$$
P\left( \left( \bigcap_{i=1}^{N} \left( \mathcal{B}_{r_{d}}(\text{supp}, \phi, \mathcal{A}_{i}) \cap \mathcal{B}_{r_{d}}^{J}(\text{supp}, \phi, \mathcal{A}_{i}) \right) \right)^{c} \right) < \varepsilon. \tag{8.3.15}$$

Hence we only have to show the convergence of

$$\mathbb{E}\left[ \exp \left( -\sum_{k=1}^{n^{*}(t)} \phi \left( \nu_{k}(r), \xi_{k}(t) \right) \right) \right]. \tag{8.3.16}$$

Observe that on the event in the indicator function in the the last line the following holds: If for any $i \in \{1, \ldots, N\}$, $\nu_{k}(r) \in [\mathcal{A}_{i}, \mathcal{A}_{i}^{c}]$ and $\xi_{k}(t) \in \text{supp}, \phi$ then also $\nu_{k}(r) \in [\mathcal{A}_{i}, \mathcal{A}_{i}^{c}]$, and vice versa. Hence (8.3.16) is equal to

$$\mathbb{E}\left[ \exp \left( -\sum_{k=1}^{n^{*}(t)} \phi \left( \nu_{k}(r), \xi_{k}(t) \right) \right) 1_{\mathcal{B}_{r_{d}}^{J}(\text{supp}, \phi, \mathcal{A}_{i})} \right]. \tag{8.3.17}$$

Now we apply again Lemma 8.5 and Lemma 8.6 to see that the quantity in (8.3.17) is equal to

$$\mathbb{E}\left[ \exp \left( -\sum_{k=1}^{n^{*}(t)} \phi \left( \nu_{k}(r), \xi_{k}(t) \right) \right) \right] + O(\varepsilon). \tag{8.3.18}$$

Introducing a conditional expectation given $\mathcal{F}_{r_{d}}$, we get (analogous to (3.16) in [4]) as $t \uparrow \infty$ that (8.3.18) is equal to

$$\lim_{t \uparrow \infty} \mathbb{E}\left[ \exp \left( -\sum_{k=1}^{n^{*}(t)} \phi \left( \nu_{k}(r), \xi_{k}(t) \right) \right) \right] = \lim_{t \uparrow \infty} \mathbb{E}\left[ \prod_{j=1}^{n^{*}(r_{d})} \mathbb{E}\left[ e^{-\phi(y_{j}(r), x_{j}(r_{d}) - m(t-r_{d}) - m(t-r_{d}) + \max_{j \neq s} |y_{j}(r_{d}) - x_{j}(r_{d})|)} | \mathcal{F}_{r_{d}} \right] \right]$$

$$= \mathbb{E}\left[ \prod_{j=1}^{n^{*}(r_{d})} \mathbb{E}\left[ e^{-\phi(y_{j}(r), x_{j}(r_{d}) - \sqrt{2r_{d}} + M)} | \mathcal{F}_{r_{d}} \right] \right],$$

where $M$ is the limit of the rescaled maximum of BBM. The last expression is completely analogous to Eq. (3.17) in [4]. Following the analysis of this expression up to Eq. (3.25) in [4], we find that (8.3.19) is equal to

$$c_{r_{d}} \mathbb{E}\left[ \exp \left( -C \sum_{j \in \mathcal{A}(r_{d})} y_{j}(r_{d}) e^{-\sqrt{2y_{j}(r_{d})}} \sum_{l=1}^{N} (1 - e^{m}) 1_{A_{l}}(\nu_{k}(r)) \left( e^{-\sqrt{2} r_{d}} - e^{-\sqrt{2} r_{d}} \right) \right) \right],$$

where $y_{j}(r_{d}) = x_{j}(r_{d}) - \sqrt{2r_{d}}$ and $\lim_{r_{d} \uparrow 1} c_{r_{d}} = 1$, and $C$ is the constant from law of the maximum of BBM. Using Lemma 8.4 (8.3.20) is in the limit as $r_{d} \uparrow \infty$ and $r \uparrow \infty$ equal to
\[
E \left[ \exp \left( -C \sum_{i=1}^{N} (1 - e^{-v}) (e^{-\sqrt{2}B_i} - e^{-\sqrt{2}B_i}) \right) (Z(A_i) - Z(A_j)) \right] \tag{8.3.21}
\]

\[
E \left[ \exp \left( \int (e^{\phi(x,y)} - 1) Z(dx) \sqrt{2}Ce^{-\sqrt{2}d}dy \right) \right].
\]

This is the Laplace functional of the process \( \widehat{\mathcal{G}}^\prime \), which proves Proposition 8.8.

To prove Theorem 8.1 we need to combine Proposition 8.8 with the results on the genealogical structure of the extremal particles of BBM obtained in [2] and the convergence of the decoration point process \( \Delta \) (see e.g. Theorem 2.3 of [2]).

**Proof (Proof of Theorem 8.1).** For \( x_k(t) \in \text{supp} \left( \mathcal{G}^{(r_d/\varepsilon)}(t) \right) \) define the process of recent relatives by

\[
\Delta_{x,t}^{(k)} = \delta_0 + \sum_{j: \tau_j^k \geq t-r} \mathcal{N}_x \Delta_{x,t}^{(k)},
\tag{8.3.22}
\]

where \( \tau_j^k \) are the branching times along the path \( s \mapsto x_k(s) \) enumerated backwards in time and \( \mathcal{N}_x \) the point measures of particles whose ancestor was born at \( \tau_j^k \). In the same way let \( \Delta_{x,t}^{(n)} \) be independent copies of \( \Delta \), which is defined as

\[
\Delta_x \equiv \lim_{t \uparrow \infty} \sum_{n(t)} \mathbb{I}_{(\hat{\xi}_t(t), \arg \max_{j \in \mathbb{N}(t)} x_j(t)) \geq t-r} \delta_{\hat{\xi}_t(t)}(t) \delta_{x}(t) - \max_{x \in \mathbb{N}(t)} x_j(t),
\tag{8.3.23}
\]

the point measure obtained from \( \Delta \) by only keeping particles that branched of the maximum after time \( t-r \) (see the backward description of \( \Delta \) in [2]). By Theorem 2.3 of [2] we have that (the labelling \( i_k \) refers to the thinned process \( \mathcal{G}^{(r_d/\varepsilon)}(t) \))

\[
\left( x_k(r_d), \sqrt{2}r_d + M_k(t - r_d), \Delta_{x,t}^{(k)} \right)_{1 \leq k \leq n(t)} \Rightarrow \left( x_j(r_d), \sqrt{2}r_d + M_j, \Delta_{x,t}^{(j)} \right)_{j \leq n(r_d)},
\tag{8.3.24}
\]

as \( t \uparrow \infty \), where \( M_j \) are independent copies of \( M \). Moreover, \( \Delta_{x,t}^{(j)} \) is independent of \( (M_{j,t})_{j \leq n(r_d)} \). Looking now at the the Laplace functional for the complete point process \( \widehat{\mathcal{G}}' \),

\[
\widetilde{\Psi}_{t}(\phi) \equiv E \left[ e^{\int \phi(x,y)\delta_{x}(dx,dy)} \right],
\tag{8.3.25}
\]

for \( \phi \) as in (8.3.9), and doing the same manipulations as in the proof of Proposition 8.8, shows that

\[
\widetilde{\Psi}_{t}(\phi) = E \left[ \exp \left( -\sum_{k=1}^{n(t)} \phi \left( \gamma_k(t), \hat{\xi}_k(t) \right) \right) \right] + O(\varepsilon).
\tag{8.3.26}
\]

Denote by \( \mathcal{G}^{t,s}_{x,r}(D) \) the event
\begin{align}
\mathcal{C}_{t,r}(D) = \forall i, j \leq n(t) \text{ with } x_i(t), x_j(t) \in D + m(t): \ d(x_i(t), x_j(t)) \notin (r, t-r). 
\end{align} 
\tag{8.3.27}

By Theorem 2.1 in \cite{3}, we know that, for each $D \subset \mathbb{R}$ compact,
\begin{align}
\lim_{r \to \infty} \sup_{t > \frac{3}{r}} \mathbb{P}(\mathcal{C}_{t,r}(D)) = 0. 
\end{align} 
\tag{8.3.28}

Hence by introducing $1 = 1_{\mathcal{C}_{t,r}(\text{supp} \phi)} + 1_{\mathcal{C}_{t,r}(\text{supp} \phi)}$ into (8.3.26), we obtain that
\begin{align}
\tilde{\Psi}(\phi) = \mathbb{E} \left[ e^{-\sum_{k=1}^{n(t)} \left( \phi(\gamma(r), \bar{x}_k(t)) + \sum_j \phi(\gamma(r), \bar{x}_j(t) + A_{j,t}^{(k,j)}) \right)} \right] + O(\varepsilon), 
\end{align} 
\tag{8.3.29}

where $A_{j,t,r_d}^{(k,j)}$ are the atoms of $\Delta_{t,r_r}^{(k)}$. Hence it suffices to show that
\begin{align}
\sum_{k=1}^{n(t)} \sum_j \delta_{(\gamma(r), \bar{x}_j(t) + A_{j,t}^{(k,j)})}
\end{align} 
\tag{8.3.30}

converges weakly when first taking the limit $t \uparrow \infty$ and then the limit $r_d \uparrow \infty$ and finally $r \uparrow \infty$. But by (8.3.24),
\begin{align}
\lim_{t \to \infty} \sum_{k=1}^{n(t)} \sum_j \delta_{(\gamma(r), \bar{x}_j(t) + A_{j,t}^{(k,j)})} = \sum_{j=1}^{n(r_d)} \sum_j \delta_{(\gamma(r), x_j(r_d) - \sqrt{2}r_d + M_j)}}
\end{align} 
\tag{8.3.31}

The limit as first $r_d$ and then $r$ tend to infinity of the process on the right-hand side exists and is equal to $\tilde{\delta}$ by 8.8 (in particular (8.3.3)). This concludes the proof of Theorem 8.1.
Chapter 9
Variable speed BBM

We have seen that BBM is somewhat related to the GREM where the covariance is a linear function of the distance. It is natural to introduce versions of BBM that have more general covariance functions \( A(x) \). This can be achieved by changing the speed (= variance) of the Brownian motion with time. Variable speed BBM was introduced by Derrida and Spohn [31] and has recently been studied recently by Fang and Zeitouni [36, 35] and others [59, 60, 63]. This entire chapter is based on joined work with Lisa Hartung [17, 18].

9.1 The construction

The general model can be constructed as follows. Let \( A : [0, 1] \to [0, 1] \) be a right-continuous increasing function. Fix a time horizon \( t \) and let

\[
\Sigma^2(u) = tA(u/t). \tag{9.1.1}
\]

Note that \( \Sigma^2 \) is almost everywhere differentiable and denote by \( \sigma^2(s) \) its derivative wherever it exists. We define Brownian motion with speed function \( \Sigma^2 \) as time change of ordinary Brownian motion on \( [0, t] \) as

\[
B^\Sigma_t = B_{\Sigma^2(t)} \tag{9.1.2}
\]

Branching Brownian motion with speed function \( \Sigma^2 \) is constructed like ordinary branching Brownian motion except that if a particle splits at some time \( s < t \), then the offspring particles perform variable speed Brownian motions with speed function \( \Sigma^2 \), i.e. their laws are independent copies \( \{B^\Sigma_r - B^\Sigma_s\}_{1 \geq r \geq s} \), all starting at the position of the parent particle at time \( s \).

We denote by \( n(s) \) the number of particles at time \( s \) and by \( \{x_i(s); 1 \leq i \leq n(s)\} \) the positions of the particles at time \( s \). If we denote by \( d(x_i(t), x_k(t)) \) the time of the most recent common ancestor of the particles \( i \) and \( k \), then a simple computation
shows that
\[ E[x_k(s)x_\ell(s)] = \Sigma^2(d(x_k(s), x_\ell(s))). \] (9.1.3)
Moreover, for different times, we have
\[ E[x_k(s)x_\ell(r)] = \Sigma^2(d(x_k(t), x_\ell(t) \land s \land r)). \] (9.1.4)

**Remark 9.1.** Strictly speaking, we are not talking about a single stochastic process, but about a family \( \{x_k(s), k \leq n(s)\}_{s \leq t} \) of processes with finite time horizon, indexed by that horizon, \( t \).

The case when \( A \) is a step function with finitely many steps corresponds to Derrida’s GREMs, with the only difference that the binary tree is replaced by a Galton-Watson tree. The case we discuss here corresponds to \( A \) being a piecewise linear function. The case when \( A \) is arbitrary has been dubbed CREM in [20] (and treated for binary regular trees). In that case the leading order of the maximum was obtained; this analysis carries over mutando mutandis to the BBM situations. Fang and Zeitouni [35] have obtained the order of the correction (namely \( t^{1/3} \)) in the case when \( A \) is strictly concave and continuous, but there are no results on the extremal process or the law of the maximum. This result has very recently strengthened by Maillard and Zeitouni [59] who proved convergence of the law of the maximum to some travelling wave and computed the next order of the correction (which is logarithmic).

### 9.2 Two-speed BBM

Understanding the piecewise linear case seems to be a prerequisite to getting the full picture. Fang and Zeitouni [35] have in this case obtained the correct order of the corrections, which is completely analogous to the GREM situation, without further information on the law of the maximum and the extremal process.

The simplest case is two-speed BBM, i.e.
\[ \sigma^2(s) = \begin{cases} \sigma_1^2, & 0 \leq s < bt \\ \sigma_2^2, & t \leq s \leq t \end{cases}, \quad 0 < b \leq 1, \] (9.2.1)
with the normalisation
\[ \sigma_1^2 b + \sigma_2^2 (1 - b) = 1. \] (9.2.2)

Note that in the case \( b = 1, \sigma_2 = \infty \) is allowed.

Fang and Zeitouni [35] showed that
\[ \max_{k \leq n(t)} x_k(t) = \begin{cases} \sqrt{2t} - \frac{1}{2\sqrt{2}} \ln t + O(1), & \text{if } \sigma_1 < \sigma_2, \\
\sqrt{2t} (\sigma_1 b + \sigma_2 (1 - b)) - \frac{3}{2\sqrt{2}} (\sigma_1 + \sigma_2) \ln t + O(1), & \text{if } \sigma_1 > \sigma_2. \end{cases} \] (9.2.3)
The second case has a simple interpretation: the maximum is achieved by adding to the maxima of BBM at time \(tb\) the maxima of their offspring at time \(t(1-b)\) later, just as in the analog case of the GREM. The first case looks like the REM, which is not surprising, as we have already seen in the GREM that correlations have no influence as long as \(A(u) < u\). But when looking at the extremal process, we will find that here things are a lot more subtle.

The main result of [17] is the following.

**Theorem 9.2 ([17]).** Let \(x_k(t)\) be branching Brownian motion with variable speed as given in (9.2.1). Assume that \(\sigma_1 < \sigma_2\) and \(b \in (0,1)\). Then

(i) \[
\lim_{t \to \infty} \mathbb{P} \left( \max_{k \leq n(t)} x_k(t) - \tilde{m}(t) \leq x \right) = \mathbb{E} \left[ e^{-C'Ye^{-\sqrt{2}t}} \right],
\]

where \(\tilde{m}(t) = \sqrt{2t} - \frac{1}{2\sqrt{2}} \ln t\), \(C'\) is a constant and \(Y\) is a random variable that is the limit of a martingale (but different from \(Z!\)).

(ii) The point process \(\tilde{E}_t \equiv \sum_{k \leq n(t)} \delta_{x_k(t) - \tilde{m}(t)} \) → \(\tilde{E}\),

as \(t \to \infty\), in law, where

\[
\tilde{E} = \sum_{k,j} \delta_{\eta_k + \sigma_2 \Lambda_{(k)}^j},
\]

where \(\eta_k\) is the \(k\)-th atom of a mixture of Poisson point process with intensity measure \(C'Ye^{-\sqrt{2}t}dx\), with \(C'\) and \(Y\) as in (i), and \(\Lambda_{(k)}^j\) are the atoms of independent and identically distributed point processes \(\Lambda_{(k)}\), which are the limits in law of

\[
\sum_{j \leq n(t)} \delta_{x_j(t)} \max_{j \leq n(t)} x_j(t),
\]

where \(x(t)\) is BBM of speed 1 conditioned on \(\max_{j \leq n(t)} x_j(t) \geq \sqrt{2}\sigma_2t\).

The picture is completed by the limiting extremal process in the case \(\sigma_1 > \sigma_2\). This result is much simpler and could be guessed from the known fact on the GREM.

**Theorem 9.3 ([17]).** Let \(x_k(t)\) be as in Theorem 9.2 but \(\sigma_2 < \sigma_1\). Again \(b \in (0,1)\). Let \(\tilde{E} \equiv \tilde{E}^{0}\) and \(\tilde{E}^{(i)}\), \(i \in \mathbb{N}\) be independent copies of the extremal process of standard branching Brownian motion. Let

\[
m(t) \equiv \sqrt{2t}(b\sigma_1 + (1-b)\sigma_2) - \frac{3}{2\sqrt{2}}(\sigma_1 + \sigma_2) \ln t - \frac{3}{2\sqrt{2}}(\sigma_1 \ln b + \sigma_2 \ln(1-b)),
\]

and set \(\tilde{E}_t \equiv \sum_{k \leq n(t)} \delta_{x_k(t) - m(t)}\).

Then
exists, and
\[ \tilde{\mathcal{E}} = \sum_{i,j} \delta \sigma_1 \epsilon_i + \sigma_2 \epsilon_j^{(i)}, \]  
(9.2.11)

where \( \epsilon_i, \epsilon_j^{(i)} \) are the atoms of the point processes \( \mathcal{E} \) and \( \mathcal{E}^{(i)} \), respectively.

We just remark on the main steps in the proof of Theorem 9.2. The first step is a
localisation of the particles that will eventually reach the top at the time of the speed
change. The following proposition says that these are in a \( \sqrt{t} \) neighbourhood of
\( \sigma_1 t \sqrt{2} b \), which is much smaller then the position \( \sigma_1 t \sqrt{2} b \) of the leading particles at
time \( tb \). Thus, the faster particles in the second half-time must make up for this. It is
not very hard to know everything about the particles after time \( tb \). The main problem
one is faced with is to control their initial distribution at this time. Fortunately, this
can be done with the help of a martingale. Define
\[ Y_s = \sum_{i=1}^{n(s)} e^{-s(1+\sigma_1^2) + \sqrt{2} \sigma_2^2 s}. \]  
(9.2.12)

When \( \sigma_1 < 1 \), one can show that \( Y \) is a uniformly integrable positive martingale
with mean value one. The martingale \( Y_s \) will take over the role of the derivative
martingale in the standard case.

We no go in more detail through the proofs of the theorems.

### 9.2.1 Position of extremal particles at time \( bt \)

The key to understanding the behaviour of the two speed BBM is to control the
positions of particle at time \( bt \) which are in the top at time \( t \). This is done using
straightforward Gaussian estimates.

**Proposition 9.4.** Let \( \sigma_1 < \sigma_2 \). For any \( d \in \mathbb{R} \) and any \( \varepsilon > 0 \), there exists a constant
\( A > 0 \) such that for all \( t \) large enough
\[ \mathbb{P} \left[ \exists j \leq n(t) \text{ s.t. } x_j(t) > \tilde{m}(t) - d \text{ and } x_j(bt) - \sqrt{2} \sigma_1^2 bt \not\in [-A \sqrt{t}, A \sqrt{t}] \right] \leq \varepsilon. \]  
(9.2.13)

**Proof.** Using a first order Chebyshev inequality we bound (9.2.13) by
\[ e^{t} \mathbb{E} \left[ \mathbb{1}_{\{\sigma_1 \sqrt{tw_1} - \sqrt{2} \sigma_1^2 bt \not\in [-A \sqrt{t}, A \sqrt{t}]\}} \mathbb{P}_{w_2} \left( \sigma_2 \sqrt{(1-b)tw_2} > \tilde{m}(t) - d - \sigma_2 \sqrt{bt w_1} \right) \right] = e^{t} \mathbb{E} \left[ \mathbb{1}_{\{w_1 - \sqrt{2} \sigma_1 \sqrt{tw_1} \not\in [-A', A']\}} \mathbb{P}_{w_2} \left( \frac{\sqrt{2} \sigma_2 \sqrt{(1-b)t} - \log t}{2 \sqrt{2} \sigma_2 \sqrt{(1-b)t}} - \frac{d}{\sigma_2 \sqrt{(1-b)t}} \right) \right] \equiv (R1) + (R2), \]  
(9.2.14)
where \( w_1, w_2 \) are independent \( \mathcal{N}(0, 1) \)-distributed, \( A' = \frac{1}{b \sigma_1} A \), \( \mathbb{P}_{w_2} \) denotes the law of the variable \( w_2 \). Introducing into the last line the identity in the form
\[
1 = \mathbb{1}_{\{2 - \sigma_1 \sqrt{bw_1} < \log t\}} + \mathbb{1}_{\{2 - \sigma_1 \sqrt{bw_1} \geq \log t\}} \quad (9.2.15)
\]
we can write it as \((R1) + (R2)\).

We first show \( \lim_{t \to \infty} (R1) = 0 \). Using the standard Gaussian tail estimate, \((R1)\) is bounded from above by
\[
e^t \mathbb{P} \left[ \sqrt{2t} - \sigma_1 \sqrt{bw_1} < \log t \right] \leq e^{t(1 - b \sigma_1^2)} + e^{t/2 \log t / b \sigma_1^2} \to 0 \quad \text{as} \ t \to \infty. \quad (9.2.16)
\]
By the same argument, \((R2)\) is smaller than
\[
e^t (2\pi)^{-1} \int_{w_1 = \sqrt{2\sigma_1 \sqrt{bw_1} \geq \log t}} 1 \frac{\sqrt{2 - \sigma_1 \sqrt{bw_1}}}{\sqrt{2t - \sigma_1 \sqrt{bw_1}}} e^{-w_1^2 / 2} \quad (9.2.17)
\]
\[
\times \exp \left( -\frac{1}{2} \left( \frac{\sqrt{2 - \sigma_1 \sqrt{bw_1} \log t / (2\sqrt{2} - \sqrt{t} - \sigma_1 \sqrt{bw_1} / \sigma_1)}^2}{1 - \sigma_1^2} \right) \right) dw_1.
\]

We change variables \( w_1 = \sqrt{2\sigma_1} \sqrt{bt} + z \). Then the integral in \((9.2.17)\) can be bounded from above by
\[
M \left( \frac{2\pi \sigma_1^2}{1 - b} \right)^{1/2} \int_{z \in [\mathcal{A}, \mathcal{A}']} e^{-\frac{1}{2} \left( \frac{2\sigma_1^2 (1 - b)}{1 - b} \right) z^2} dz, \quad (9.2.18)
\]
where \( M \) is some positive constant. \((9.2.18)\) can be made as small as desired by taking \( A \) (and thus \( A' \)) sufficiently large. \( \square \)

**Remark 9.5.** The point here is that since \( \sigma_1^2 < \sigma_1 \), these particles are way below \( \max_{k \leq n(ht)} x_k(bt) \), which is near \( \sqrt{2\sigma_1} \sigma_1 \). The offspring of these particles that want to be top at time will have to race much faster (at speed \( \sqrt{2\sigma_1} \), rather than just \( \sqrt{2\sigma_1} \)) than normal. Fortunately, there are lots of particles to choose from. We will have to control precisely how many.

We need a slightly finer control on the path of the extremal particle until the time of speed change. To this end we define two sets on the space of paths, \( \mathcal{X} \) : \( \mathbb{R}_+ \to \mathbb{R} \). The first controls that the position of the path is in a certain tube up to time \( s \) and the second the position of the particle at time \( s \). We define the following events.
\[
\mathcal{F}_{s, r} = \{ X | \forall 0 \leq q \leq s, |X(q) - \frac{q}{s} X(s)| \leq (|q \wedge (s - q)|) \vee r \}
\]
\[
\mathcal{O}_{s, A, Y} = \{ X | X(s) - \sqrt{2\sigma_1^2} s \in [-As', +As'] \} \quad (9.2.19)
\]
Recall that the ancestral path form 0 to \( x_k(s) \) can be written as \( x_k(q) = \frac{q}{s} x_k(s) + \tilde{z}_k(s) \), where \( \tilde{z}_k \) is a Brownian bridge from 0 to 0 in time \( s \), independent of \( x_k(s) \). We need the following simple fact about Brownian bridges.
Lemma 9.6. Let \( \mathcal{B}(q) \) be a Brownian bridge starting in zero and ending in zero at time \( s \). Then for all \( \gamma > 1/2 \), the following is true. For all \( \epsilon > 0 \) there exists \( r \) such that
\[
\lim_{\epsilon \to 0} \mathbb{P}\left(|\mathcal{B}(q)| < (|q \wedge (s-q)|) \mathbb{V} r\right) \geq 0 \quad \forall q \leq s > 1 - \epsilon. \tag{9.2.20}
\]

Proposition 9.7. Let \( \sigma_1 < \sigma_2 \). For any \( d \in \mathbb{R}, A > 0, \gamma > 1/2 \) and any \( \epsilon > 0 \), there exists constants \( B > 0 \) such that, for all \( t \) large enough,
\[
\mathbb{P}\left(\exists j \leq n(t) : x_j(t) > \tilde{m}(t) - d \wedge x_j \in \mathcal{B}_{bt,A,1/2} \wedge x_j \notin \mathcal{B}_{bt,A,2}\gamma\right) \leq \epsilon. \tag{9.2.21}
\]

Proof. For \( B \) and \( t \) sufficiently large the probability in (9.2.21) is bounded from above by
\[
\mathbb{P}\left(\exists j \leq n(t) : x_j(t) > \tilde{m}(t) - d \wedge x_j \in \mathcal{B}_{bt,A,1/2} \wedge x_j \notin \mathcal{B}_{bt}\right) \tag{9.2.22}
\]
Let \( w_1 \) and \( w_2 \) be independent \( \mathcal{N}(0,1) \)-distributed random variables and \( \mathcal{B} \) a Brownian bridge starting in zero and ending in zero at time \( bt \). Using a first moment method together with the independence of the Brownian bridge from its endpoint, one sees that (9.2.22) is bounded from above by
\[
e^{-\epsilon} \mathbb{E}\left[\|\sigma_1 \sqrt{w_1 - \sqrt{\mathcal{B}_{bt,A,\gamma}}}\|\mathcal{P}_{w_2}\left(\sigma_2 \sqrt{(1-b)w_2} > \tilde{m}(t) - d - \sigma_1 \sqrt{btw_1}\right)\right]
\times \mathbb{P}\left(\mathcal{B} \notin \mathcal{B}_{bt}\right) < \epsilon, \tag{9.2.23}
\]
where the last bound follows from Lemma 9.6 (with \( \epsilon \) replaced by \( \epsilon/M \)) and the bound (9.2.18) obtained in the proof of Proposition 9.4. \( \square \)

Proposition 9.8. Let \( \sigma_1 < \sigma_2 \). For any \( A, B > 0, \gamma > 1/2 \) and any \( \epsilon > 0 \), there exists a constant \( r > 0 \) such that for all \( t \) large enough
\[
\mathbb{P}\left(\exists j \leq n(t) : x_j(t) > \tilde{m}(t) - d \wedge x_j \in \mathcal{B}_{bt,A,1/2} \wedge x_j \notin \mathcal{B}_{bt,A,2}\gamma \wedge \mathcal{B}_{bt(A-\gamma)} \right) \leq \epsilon. \tag{9.2.24}
\]

Proof. The proof of this proposition is essentially identical to the proof of Proposition 9.7. \( \square \)

### 9.2.2 Convergence of the McKean martingale

In this section we pick up the idea of [56] and consider a suitable convergent martingale for the time inhomogeneous BBM with \( \sigma_1 < \sigma_2 \). Let \( x_i(t), 1 \leq i \leq n(t) \) be the particles of a BBM where the Brownian motions have variance \( \sigma_i^2 \) with \( \sigma_i^2 < 1 \). Define
9.2 Two-speed BBM

\[ Y_s = \sum_{i=1}^{n(s)} e^{-s(1+\sigma_i^2)+\sqrt{s}z_i(s)}. \]  
(9.2.25)

This turns out to be a uniformly integrable martingale that converges almost surely to a positive limit \( Y \).

**Theorem 9.9.** If \( \sigma_1 < 1 \), \( Y_s \) is a uniformly integrable martingale with \( \mathbb{E}[Y_s] = 1 \). Hence the limit \( \lim_{s \to \infty} Y_s \) exists almost surely and in \( L^1 \), is finite and strictly positive.

**Remark 9.10.** Note further that if \( \sigma_1 = 0 \), then \( Y_i = e^{-s}n(t) \) which is well known to converge to an exponential random variable.

**Proof.** Clearly,
\[ \mathbb{E}[Y_s] = e^{s} \mathbb{E}\left[e^{-s(1+\sigma_i^2)+\sqrt{s}z_i(s)}\right] = 1. \]  
(9.2.26)

Next we show that \( Y_s \) is a martingale. Let \( 0 < r < s \). Then
\[ \mathbb{E}[Y_s | \mathcal{F}_r] = \sum_{i=1}^{n(r)} \mathbb{E}\left[ \sum_{j=1}^{n_i(s-r)} e^{-s(1+\sigma_i^2)+\sqrt{s}z_i(s-r)+x_i(r)} | \mathcal{F}_r \right], \]  
(9.2.27)

where for \( 1 \leq i \leq r \), \( \left\{ x_i'(s-r), 1 \leq j \leq n_i(s-r) \right\} \) are the particles of independent BBM’s with variance \( \sigma_i^2 \) at time \( s-r \). (9.2.27) is equal to
\[ \sum_{i=1}^{n(r)} e^{-r(1+\sigma_i^2)+\sqrt{r}z_i(r)} = Y_r, \]  
(9.2.28)

as desired.

It remains to show that \( Y_s \) is uniformly integrable. Recall that \( x_k(r) \) is the ancestor of \( x_k(s) \) at time \( r \leq s \) and write \( x_k \) for the entire ancestral path of \( x_k(s) \). Define the truncated variable
\[ Y_s^A = \sum_{i=1}^{n(s)} e^{-s(1+\sigma_i^2)+\sqrt{s}z_i(s)}1_{\{x_i(t) \cap A = t, x_i(t) \cap \mathcal{F}_s \}}. \]  
(9.2.29)

First \( Y_s - Y_s^A \geq 0 \), and a simple computation shows that
\[
\mathbb{E} \left[ Y_s - Y_s^A \right] \leq \mathbb{E} \left[ \sum_{i=1}^{n(s)} e^{-s(1+\sigma_i^2)+\sqrt{s}z_i(s)}1_{\{x_i \cap \mathcal{F}_A \}} \right] \\
+ \mathbb{E} \left[ \sum_{i=1}^{n(s)} e^{-s(1+\sigma_i^2)+\sqrt{s}z_i(s)}1_{\{x_i \not\in \mathcal{F}_s \}} \right] \\
\leq e^{\epsilon} \int_{-\infty}^{\infty} e^{-s(1+\sigma_i^2)+\sqrt{s}z_i(s)}1_{\{z_i \geq \sqrt{s}\sqrt{\sigma_i}1_{\not \in [-A,A]}\} e^{-s^2/2} \frac{dz}{\sqrt{2\pi}} + \epsilon} \\
= \int_{|z| > A} e^{-z^2/2} \frac{dz}{\sqrt{2\pi}} + \epsilon, \]  
(9.2.30)
where we used that \( x_i(s) \) and the bridge \( x_i(r) - x_i(s) \) are independent and that the event \( \{ x_i \notin \mathcal{F}_r \} \) depends only on this bridge and by Lemma 9.6 has probability less than \( \epsilon \). The bound in (9.2.30) can be made as small as desired by taking \( A \) and \( T \) to infinity. The key point is that the second moment of \( Y_s^A \) is uniformly bounded in \( s \).

\[
\mathbb{E} \left[ (Y_s^A)^2 \right] = \mathbb{E} \left[ \left( \sum_{k=1}^{n(s)} e^{-i(1+\sigma_k^2)x+\sqrt{2s}z_k} \mathbb{I}_{\{ x_k \in \mathcal{B}_{s,1} \cap \mathcal{F}_r \}} \right)^2 \right] \equiv (T1) + (T2),
\]

(9.2.31)

where

\[
(T1) = \mathbb{E} \left[ \sum_{k=1}^{n(s)} e^{-i(1+\sigma_k^2)x+\sqrt{2s}z_k} \mathbb{I}_{\{ x_k \in \mathcal{B}_{s,1} \cap \mathcal{F}_r \}} \right]
\]

\[
(T2) = \mathbb{E} \left[ \sum_{i=1}^{n(s)} e^{-i(1+\sigma_i^2)x+\sqrt{2s}z_i} \mathbb{I}_{\{ x_i \in \mathcal{B}_{s,1} \cap \mathcal{F}_r \}} \right]
\]

(9.2.32)

We start by controlling \((T1)\).

\[
(T1) \leq \frac{e^{(s-2\sigma(1+\sigma_i^2))}}{\sqrt{2\pi}} \int_{-A/|\sigma_1|}^{A/|\sigma_1|} e^{-x^2/2} dx
\]

\[
\leq \frac{e^{-(1-\sigma_i^2)^2}}{\sqrt{2\pi}} \int_{-A/|\sigma_1|}^{A/|\sigma_1|} e^{-x^2/2} dx \leq e^{-(1-\sigma_i^2)^2} e^{A^2/2} \rightarrow 0 \quad \text{as } s \rightarrow \infty. (9.2.33)
\]

Note that here we used the condition that \( x_i \in \mathcal{B}_{s,1/2} \), otherwise this contribution would diverge. Now we control \((T2)\). Using a second moment bound and dropping the useless parts of the conditions on the Brownian bridges

\[
(T2) \leq Ke^s \int_0^s e^{s-q} \int_{-I_1(q,s)}^{I_1(q,s)} e^{-q} \frac{e^{2\sigma^2(x-y)}}{\sigma_1 \sqrt{2\pi(x-y)^2}} \frac{dy}{\sigma_1 \sqrt{2\pi(x-y)^2}} + \int_{s-q}^{s} e^{-q} \frac{e^{2\sigma^2(x-y)}}{\sigma_1 \sqrt{2\pi(x-y)^2}} \frac{dy}{\sigma_1 \sqrt{2\pi(x-y)^2}}
\]

(9.2.34)

where \( K = \sum_{k=1}^{\infty} p_{k,k}(k-1) \) and \( I_1(q,s) = Aq/\sqrt{s} + ((q \wedge (s-q)) \vee r)^r \). Moreover we change variables \( x = z + \sqrt{2\sigma^2}q \) and obtain

\[
Ke^s \int_0^s e^{s-q} \int_{-I_1(q,s)}^{I_1(q,s)} e^{-q} \frac{e^{2\sigma^2(x-y)}}{\sigma_1 \sqrt{2\pi(x-y)^2}} \frac{dy}{\sigma_1 \sqrt{2\pi(x-y)^2}} + \int_{s-q}^{s} e^{-q} \frac{e^{2\sigma^2(x-y)}}{\sigma_1 \sqrt{2\pi(x-y)^2}} \frac{dy}{\sigma_1 \sqrt{2\pi(x-y)^2}}
\]

(9.2.35)
9.2 Two-speed BBM

Now we change variables \( w = \frac{s}{\sigma_1 \sqrt{q}} - \sqrt{2} \sigma_1 \sqrt{s - q} \). (9.2.35) is equal to

\[
K \int_0^s e^{-q(1-2\sigma_1^2)} \int_{-I_1(q,s)}^{+I_1(q,s)} e^{2\sqrt{2} \pi} \left( \frac{i+\sqrt{2} \sigma_1 \sqrt{q}}{\sigma_1 \sqrt{q-\sigma}} \right) e^{-w^2/2} \frac{dw}{\sqrt{2\pi}} e^{-\frac{(i+\sqrt{2} \sigma_1 \sqrt{q})^2}{2\sigma_1 \sqrt{q}}} dq, \tag{9.2.36}
\]

Now the integral with respect to \( w \) is bounded by 1. Hence (9.2.36) is bounded from above by

\[
K \int_0^s e^{-q(1-2\sigma_1^2)} \int_{-I_1(q,s)}^{+I_1(q,s)} e^{-\frac{(i+\sqrt{2} \sigma_1 \sqrt{q})^2}{2\sigma_1 \sqrt{q}}} dq. \tag{9.2.37}
\]

We split the integral over \( q \) into the three parts \( R_1, R_2, \) and \( R_3 \) according to the integration from 0 to \( r \), \( r \) to \( s-r \), and \( s-r \) to \( r \), respectively. Then

\[
R_2 \leq K \int_r^{s-r} e^{-q(1-2\sigma_1^2)} e^{-\frac{1}{2} \left( I_1(q,s)/\sigma_1 \sqrt{q - \sqrt{2} \sigma_1 \sqrt{q}} \right)^2} dq. \tag{9.2.38}
\]

This is bounded by

\[
K \int_r^{s-r} e^{-(1-\sigma_1^2)q+O(q^2)} dq \leq \frac{C}{1-\sigma_1^2} e^{-(1-\sigma_1^2)r}, \tag{9.2.39}
\]

For \( R_1 \) the integral over \( z \) can only be bounded by one. This gives

\[
R_1 \leq K \int_0^r e^{2(\sigma_1^2-1)q} dq = D_1(r), \tag{9.2.40}
\]

\( R_3 \) can be treated the same way as \( R_2 \) and we get

\[
R_3 \leq K \int_{s-r}^{s} e^{-(1-\sigma_1^2)q+O(r)} dq \leq \frac{K}{1-\sigma_1^2} e^{-(1-\sigma_1^2)(s-r)+O(r)} \to 0 \quad \text{as} \quad s \to \infty. \tag{9.2.41}
\]

Putting all three estimates together, we see that \( \sup_t E \left[ (Y_s^A)^2 \right] \leq D_2(\tau) \). From this it follows that \( Y_s \) is uniformly integrable. Namely,

\[
E[Y_s 1_{Y_s \geq z}] = E[Y_s^A 1_{Y_s \geq z}] + E[(Y_s - Y_s^A) 1_{Y_s \geq z}]
\]

\[
= E[Y_s^A 1_{Y_s \geq z/2}] + E \left[ Y_s^A \left( 1_{Y_s \geq z - 1_{Y_s \geq z/2}} \right) \right] + E[(Y_s - Y_s^A) 1_{Y_s \geq z}]. \tag{9.2.42}
\]

For the first term we have

\[
E \left[ Y_s^A 1_{Y_s \geq z/2} \right] \leq \frac{2}{z} E \left[ (Y_s^A)^2 \right] \leq \frac{2}{z} D_2(\tau). \tag{9.2.43}
\]

For the second, we have
\[
\mathbb{E} \left[ Y_s^A \left( 1_{Y_s > z} - 1_{Y_s^A > z/2} \right) \right] \leq \mathbb{E} \left[ Y_s^A 1_{Y_s - Y_s^A \geq z/2} 1_{Y_s^A \leq z/2} \right] \\
\leq \frac{1}{2} \mathbb{P} \left[ (Y_s - Y_s^A) > z/2 \right] \leq \mathbb{E} \left[ Y_s - Y_s^A \right].
\]

The last term in (9.2.42) is also bounded by \( \mathbb{E} \left[ Y_s - Y_s^A \right] \). Choosing now \( A \) and \( r \) such that \( \mathbb{E} \left[ Y_s - Y_s^A \right] \leq \epsilon / 3 \), and then \( z \) so large that \( \frac{z}{2} D_2(r) \leq \epsilon / 3 \), we obtain that \( \mathbb{E}[Y_s 1_{Y_s > z}] \leq \epsilon \), for large enough \( z \), uniformly in \( s \). Thus \( Y_s \) is uniformly integrable, which we wanted to show. The remainder of the assertion of the theorem is now standard theory. □

We will also need to control the processes \( \tilde{Y}_{s,A} = \sum_{i=1}^{n(s)} e^{-\left(1+\sigma_i^2\right)x+i\sqrt{\sigma_i}(s)} 1_{x_i \in [A,A+Y]}. \)

**Lemma 9.11.** The family of random variables \( \tilde{Y}_{s,A}, s,A \in [0,1], r > 1/2 \) is uniformly integrable and converges, as \( s \uparrow \infty \) and \( A \uparrow \infty \), to \( Y \), both in probability and in \( L^1 \).

**Proof.** The proof of uniform integrability is a rerun of the proof of Theorem 9.9 noting that the bounds on the truncated second moments are uniform in \( A \). Moreover, the same computation as in Eq. (9.2.30) shows that \( \mathbb{E}[Y_s - \tilde{Y}_{s,A}] \leq \epsilon \), uniformly in \( s \), for \( A \) large enough. Therefore,

\[
\lim_{s \uparrow \infty} \limsup_{A \uparrow \infty} \mathbb{E}[Y_s - \tilde{Y}_{s,A}] = 0,
\]

which implies that \( Y_s - \tilde{Y}_{s,A} \) converges to zero in probability. Since \( Y_s \) converges to \( Y \) almost surely, we arrive at the second assertion of the lemma. □

### 9.2.3 Asymptotic behaviour of BBM

We now need to control how the particles in the second epoch of high variance make it up to the top. This requires control on the asymptotics of the F-KPP equation even further ahead of the travelling wave than was done in Chapter 7. More precisely, we need to control solutions \( u(t,x + \sqrt{2t}) \) for values \( x = at, a > 0 \).

The following proposition is an extension of Lemma 7.3 for these values of \( x \).

**Lemma 9.12.** Let \( u \) be a solution to the F-KPP equation with initial data satisfying

(i) \( 0 \leq u(0,x) \leq 1 \);

(ii) for some \( h > 0 \), \( \limsup_{t \to \infty} \frac{1}{7} \log \int_t^{t+h} u(0,y)dy \leq -\sqrt{2} \);

(iii) for some \( v > 0 \), \( M > 0 \), \( N > 0 \), it holds that \( \int_x^{x+N} u(0,y)dy > v \) for all \( x \leq -M \);

(iv) moreover, \( \int_0^x u(0,y)e^{2y}dy \leq \infty \).

Then we have for \( x = at + o(t) \)

\[
\lim_{t \to \infty} e^{\sqrt{2t}e^{t/2}}u(t,x + \sqrt{2t}) = C(a),
\]

(9.2.46)
where $C(a)$ is a strictly positive constant. The convergence is uniform for $a$ in compact intervals.

**Proof.** Recall the definition of the function $\psi(r,t,x+\sqrt{2t})$ from (6.1.9). As in the proof of Lemma 7.3, the claimed result follows from the next lemma that provides the asymptotics of $\psi$.

**Lemma 9.13.** For $x = at + o(t)$ we have, under the assumptions of Proposition 9.12

$$\lim_{t \to \infty} e^{\sqrt{2t}e^{x^2/2t}1/2} \psi(r,t,x+\sqrt{2t}) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-a^2t/2}u(r,y+\sqrt{2t})e^{(\sqrt{2}+a)y} \left(1 - e^{-2ay}\right) dy \equiv C(r,a).$$

The convergence is uniform for $a$ in a compact set.

The proofs of this lemma is fairly analogous to the corresponding estimate in Chapter 7 and we do not give the details. See again [17]. The proof of Lemma 9.12 is now concluded as in the proof of Proposition 7.1. □

We also need the following variant of Lemma 7.6.

**Lemma 9.14.** Let $u$ be a solution of the F-KPP equation with initial data satisfying Assumptions (i)-(iv) of Proposition 9.12. Let

$$C(a) = \lim_{t \to \infty} \frac{1}{\sqrt{2\pi}} \int_0^\infty u(t,z+\sqrt{2t})e^{(\sqrt{2}+a)z-a^2t/2} dz.$$ (9.2.48)

Then for any $x \in \mathbb{R}$:

$$\lim_{t \to \infty} \frac{1}{\sqrt{2\pi}} \int_0^\infty u(t,x+z+\sqrt{2t})e^{(\sqrt{2}+a)z-a^2t/2} dz = C(a)e^{-(\sqrt{2}+a)x}.$$ (9.2.49)

Moreover, for any bounded continuous function $h(x)$, that is zero for $x$ small enough

$$\lim_{t \to \infty} \int_0^\infty \mathbb{E} \left[h(\bar{x}(t) - \sqrt{2t})\right] \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{2}+a)y-a^2t/2} dy = C(a) \int_\mathbb{R} h(z)(\sqrt{2}+a)e^{-(\sqrt{2}+a)z} dz.$$ (9.2.50)

where $\{\bar{x}_i(t), i \leq n(t)\}$ are the particles of a standard BBM with variance 1. Here $C(a)$ is the constant from (9.2.48) for $u$ satisfying the initial condition $\mathbb{1}_{x \leq 0}$.

We skip the proof that is completely analogous to that of Lemma 7.6.

### 9.2.4 Convergence of the maximum

We first show the convergence of the law the of the maximum. The following theorem is more precise then the statement in Theorem 9.2.
Theorem 9.15. Let \( \{x_k(t), 1 \leq k \leq n(t)\} \) be the particles of two speed BBM with \( \sigma_1 < \sigma_2 \). Set \( a = \sqrt{2}(\sigma_2 - 1) \). Then

\[
\lim_{t \to \infty} \mathbb{P} \left[ \max_{1 \leq k \leq n(t)} x_k(t) - \tilde{m}(t) \leq y \right] = \mathbb{E} \left[ \exp \left( -\sigma_2 C(a) Y e^{-\sqrt{2}y} \right) \right].
\] (9.2.51)

\( Y \) is the limit of the McKean martingale from the last section, and \( 0 < C(a) < \infty \) is the positive constant given by

\[
C(a) = \lim_{r \to \infty} \int_0^\infty e^{-a^2r/2} \mathbb{P} \left[ \max_{k \leq n(r)} \tilde{x}_k(r) > z + \sqrt{2}r \right] e^{(\sqrt{2}a)z} \left( 1 - e^{-2ac} \right) dz,
\] (9.2.52)

where \( \{\tilde{x}_k(t), k \leq n(t)\} \) are the particles of a standard BBM.

Proof. Denote by \( \{x_i(br), 1 \leq i \leq n(br)\} \) the particles of a BBM with variance \( \sigma_1 \) at time \( bt \) and by \( \mathcal{G}_bt \) the \( \sigma \)-algebra generated this BBM. Moreover, for \( 1 \leq i \leq n(bt) \), let \( \{x'_j((1-b)t), 1 \leq j \leq n_i((1-b)t)\} \) denote the particles of independent BBM with variance \( \sigma_2 \) at time \( (1-b)t \).

By second moment estimates one can easily show that the order of the maximum is \( \tilde{m}(t) \), i.e. for any \( \varepsilon > 0 \), there exists \( d < \infty \), such that

\[
\mathbb{P} \left[ \max_{1 \leq k \leq n(t)} x_k(t) - \tilde{m}(t) \leq -d \right] \leq \varepsilon / 2.
\] (9.2.53)

Therefore,

\[
\mathbb{P} \left[ -d \leq \max_{1 \leq k \leq n(t)} x_k(t) - \tilde{m}(t) \leq y \right] \leq \mathbb{P} \left[ \max_{1 \leq k \leq n(t)} x_k(t) - \tilde{m}(t) \leq y \right] \leq \mathbb{P} \left[ -d \leq \max_{1 \leq k \leq n(t)} x_k(t) - \tilde{m}(t) \leq y \right] + \varepsilon / 2
\] (9.2.54)

For details see [35]. On the other hand, by Proposition 9.4, all the particles that may contribute to the event in question are localised at time \( bt \) in the narrow window described by \( \Theta_{M,A,1/2} \). Thus we obtain
\[ \mathbb{P} \left[ \max_{1 \leq k \leq n(t)} x_k(t) - \bar{m}(t) \leq y \right] \]
\[ = \mathbb{P} \left[ \max_{1 \leq j \leq n((1 - b)t)} \max_{1 \leq i \leq n((1 - b)t)} x_i((1 - b)t) + x_i^-((1 - b)t) - \bar{m}(t) \leq y \right] \]
\[ = E \left[ \prod_{1 \leq j \leq n((1 - b)t)} \mathbb{P} \left[ \max_{1 \leq i \leq n((1 - b)t)} x_i^-((1 - b)t) \leq \bar{m}(t) - x_i((1 - b)t) + y \mid \tilde{\mathcal{F}}_{bt} \right] \right] \]
\[ \leq E \left[ \prod_{1 \leq j \leq n((1 - b)t)} \mathbb{P} \left[ \max_{1 \leq i \leq n((1 - b)t)} \sigma_2^{-1} x_i^-((1 - b)t) \leq \sigma_2^{-1} (\bar{m}(t) - x_i((1 - b)t) + y) \mid \tilde{\mathcal{F}}_{bt} \right] \right] \]
\[ + \epsilon. \quad (9.2.55) \]

The corresponding lower bound holds without the \( \epsilon \). We write the probability in (9.2.55) as

\[ 1 - \mathbb{P} \left[ \max_{1 \leq j \leq n((1 - b)t)} x_j^-((1 - b)t) > \sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y) \mid \tilde{\mathcal{F}}_{bt} \right] \]
\[ = 1 - u \left( (1 - b)t, \sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y) \right), \quad (9.2.56) \]

where \( x_j^-((1 - b)t) \) are the particles of a standard BBM and \( u \) is the solution of the F-KPP equation with Heaviside initial conditions. By Proposition 9.12 and setting

\[ C_j(x) \equiv e^{\sqrt{2}x + x^2/2t^{1/2}u(t, x + \sqrt{2}t)}, \quad (9.2.57) \]

we have that

\[ u \left( (1 - b)t, \sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y) \right) \]
\[ = C_{(1 - b)t} \left( (1 - b)t, \sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y) \right) \]
\[ \times e^{-\sqrt{2} \left( \frac{\sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y)}{\sqrt{2} t} - \sqrt{2}(1 - b) \right)} e^{-\frac{1}{\sigma_2} \left( \frac{\sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y)}{\sqrt{2} t} - \sqrt{2}(1 - b) \right)^2} ((1 - b)t)^{-1/2} \]

Now all the \( x_j((1 - b)t) \) that appear are of the form \( x_j((1 - b)t) = \sqrt{2} \sigma_2^2 b t + O(\sqrt{t}) \), so that

\[ C_{(1 - b)t} \left( (1 - b)t, \sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y) - \sqrt{2}(1 - b) t \right) = C_{(1 - b)t} (a(1 - b)t + O(\sqrt{t})), \quad (9.2.59) \]

where (using (9.2.2))

\[ a \equiv \frac{1}{1 - b} \left( \sqrt{2} - \sqrt{2} \sigma_2^2 b \right) = \sqrt{2}(\sigma_2 - 1). \quad (9.2.60) \]

By Proposition 9.12

\[ \lim_{t \to \infty} C_{(1 - b)t} \left( (1 - b)t, \sigma_2^{-1} (\bar{m}(t) - x_j((1 - b)t) + y) - \sqrt{2}(1 - b) t \right) = C(a), \quad (9.2.61) \]
with uniform convergence for all \(i\) appearing in (9.2.55) and \(C(a)\) is the constant given by (9.2.52). After a little algebra we can rewrite the expectation in (9.2.55) as

\[
\mathbb{E}\left[ \prod_{1 \leq i \leq n(b) \atop x_i \in \Theta_{b,A,1/2}} \left\{ 1 - C(a)((1-b)t)^{-1/2} e^{-(1-b)t - \frac{(\hat{m}(t) + y - x_i(bt))^2}{2(1-b)\sigma_i^2}} (1 + o(1)) \right\} \right]. \tag{9.2.62}
\]

Using that \(x_i(bt) - \sqrt{2}\sigma_i^2 tb \in [-A\sqrt{t}, A\sqrt{t}]\) we have the uniform bounds

\[
\exp\left( (1-b)t - \frac{(\hat{m}(t) + y - x_i(bt))^2}{2(1-b)\sigma_i^2} \right) \leq \exp\left( (1 - \sigma_i^2) bt + \log t + A\sqrt{t} \right), \tag{9.2.63}
\]

which is exponentially small in \(t\) since \(\sigma_i^2 > 1\).

Hence (9.2.62) is equal to

\[
\mathbb{E}\left[ \prod_{1 \leq i \leq n(b) \atop x_i \in \Theta_{b,A,1/2}} \exp\left( -C(a)((1-b)t)^{-1/2} e^{-(1-b)t - \frac{(\hat{m}(t) + y - x_i(bt))^2}{2(1-b)\sigma_i^2}} (1 + o(1)) \right) \right]. \tag{9.2.64}
\]

Expanding the square in the exponent in the last line and keeping only the relevant terms yields

\[
(1-b)t - \frac{(\hat{m}(t) + y - x_i(bt))^2}{2(1-b)\sigma_i^2} = (1-b)t - \sqrt{2}y - t\sigma_i^2(1-b) + 2\sigma_i^2 bt + \sqrt{2}x_i(bt) - \frac{1}{2}\ln t
\]

\[
+ \frac{(\sqrt{2}\sigma_i^2 b - x_i(bt))^2}{2(1-b)\sigma_i^2 t} + o(1)
\]

\[
= -\sqrt{2}y - bt(1 + \sigma_i^2) + \sqrt{2}x_i(bt) + \frac{1}{2}\ln t - \frac{(\sqrt{2}\sigma_i^2 b - x_i(bt))^2}{2(1-b)\sigma_i^2 t} + o(1). \tag{9.2.65}
\]

The terms up to the last one would nicely combine to produce the McKean martingale as coefficient of \(C(a)\). Namely,

\[
\prod_{1 \leq i \leq n(b) \atop x_i \in \Theta_{b,A,1/2}} \exp\left( -C(a)((1-b)t)^{-1/2} e^{-\sqrt{2}y - bt(1 + \sigma_i^2) + \sqrt{2}x_i(bt) + \frac{1}{2}\ln t} \right)
\]

\[
= \exp\left( -\frac{C(a)}{\sqrt{1-b}} e^{-\sqrt{2}y} \sum_{1 \leq i \leq n(b) \atop x_i \in \Theta_{b,A,1/2}} e^{-bt(1 + \sigma_i^2) + \sqrt{2}x_i(bt)} \right). \tag{9.2.66}
\]
However, the last terms are of order one and cannot be neglected. To deal with them, we split the process at time $b\sqrt{t}$. We write somewhat abusively $x_i(bt) = x_i(b\sqrt{t}) + x_i^{(i)}(b(t - \sqrt{t}))$, where we understand that $x_i(b\sqrt{t})$ is the ancestor at time $b\sqrt{t}$ of the particle that at time $t$ is labeled $i$ if we think backwards from time $t$, while the labels of the particles at time $b\sqrt{t}$ run only over the different ones, i.e. up to $n(b\sqrt{t})$, if we think in the forward direction. No confusion should occur if this is kept in mind.

Using Proposition 9.7 and Proposition 9.8 we can further localise the path of the particle. Recall the definition of $\mathcal{G}_{s,A,t}$ and $\mathcal{J}_{t,s}$, we rewrite (9.2.64), up to a term of order $\epsilon$, as

$$
\mathbb{E}\left[ \prod_{1 \leq i \leq n(b\sqrt{t})} \left( \prod_{x_i \in \mathcal{G}_{b\sqrt{t},b,T}} \mathbb{E}_{b\sqrt{t},b,T} \exp \left( -C(a)((1-b)t)^{1/2} \right) \times \exp \left( (1-b)t - \frac{\bar{m}(t) + x_i(b\sqrt{t}) - x_i^{(i)}(b(t - \sqrt{t}))}{2(1-b)\sigma^2 t} \right) \right) \right].
$$

Using that $x_i(b\sqrt{t}) + x_i^{(i)}(b(t - \sqrt{t})) - \sqrt{2}\sigma^2 tb \in [-A\sqrt{t},A\sqrt{t}]$ and $\bar{m}(t) = \sqrt{2t} - \frac{1}{2\sqrt{2}} \log t$, we can re-write the terms multiplying $C(a)$ in (9.2.67) as

$$
\exp \left( (1 - \sqrt{2}\sigma^2 t) \right) - \sqrt{2}\sigma^2 tb \in [-A\sqrt{t},A\sqrt{t}]
$$

$$
\equiv E(x_i, x_i^{(i)}) = E(x_i(b\sqrt{t}), x_i^{(i)}(b(t - \sqrt{t})) = E(x_i(b\sqrt{t}), x_i(bt) - x_i(b\sqrt{t})).
$$

Now (9.2.67) takes the form

$$
\mathbb{E}\left[ \prod_{1 \leq i \leq n(b\sqrt{t})} \left( \prod_{x_i \in \mathcal{G}_{b\sqrt{t},b,T}} \mathbb{E}_{b\sqrt{t},b,T} \exp \left\{ \sum_{1 \leq i \leq n(b\sqrt{t})} C(a)E(x_i, x_i^{(i)})(1 + o(1)) \right\} \right) \right].
$$

The idea is that the exponential in the conditional expectation can be replaced by its linear approximation, and that the linear term can be nicely computed. With

$$
\chi \equiv \sum_{1 \leq i \leq n(b\sqrt{t})} E(x_i, x_i^{(i)}),
$$

and performing some fairly straightforward algebra, we see that
\[ \mathbb{E}[x | \tilde{h}_{b\sqrt{t}}] \leq e^{\lambda(t - \sqrt{t}) - \sqrt{t}y} \int_{K_t - A\sqrt{t}}^{K_t + A\sqrt{t}} e^{\sqrt{t}(z + x_t(b\sqrt{t}))/\sigma^2 \sqrt{t}(1 - b)} e^{-\frac{\lambda^2 z^2}{2\sigma^2}(1 - b)t - \sqrt{2\pi \sigma^2 b(1 - b)t}} dz = e^{-(1 + \sigma^2_t) b\sqrt{t} + \sqrt{t}z_t(b\sqrt{t}) - \frac{1}{2} \log(1 - b) - \sqrt{t}y} \left( \frac{\sigma^2_t(1 - b)}{1 - \sigma^2_t/b \sqrt{t}} \right)^{1/2} \int_{-B\sqrt{t}}^{B\sqrt{t}} e^{-w^2/2t} \frac{dw}{\sqrt{2\pi t}} \]

\[ = e^{-(1 + \sigma^2_t) b\sqrt{t} + \sqrt{t}z_t(b\sqrt{t}) - \sqrt{t}y} \sigma_2(1 + o(1)), \quad (9.2.71) \]

where \( o(1) \leq O(t^{-1}) \) and \( B = A \left( \sqrt{b(1 - b/\sqrt{t})} \right) \). Note that the inequality comes from the fact that we dropped the tube conditions. However, by the fact that the Brownian bridge from its endpoint and that the bridge verifies the tube condition with probability at least \( 1 - \varepsilon \) (see Lemma 9.6), it follows that the right hand side of (9.2.71) multiplied by an additional factor \( (1 - \varepsilon) \) is a lower bound. This is what we want. To justify the replacement of the exponential by the linear approximation, due to the the inequalities

\[ 1 - x \leq e^{-x} \leq 1 - x + \frac{1}{2} x^2, \quad x > 0, \quad (9.2.72) \]

we just need to control the second moment. But

\[ \mathbb{E}[x^2 | \tilde{h}_{b\sqrt{t}}] \leq e^{-2(1 + \sigma^2_t) b\sqrt{t} + 2\sqrt{t}z_t(b\sqrt{t}) - 2\sqrt{t}y} \mathbb{E} \left[ Y_{b\sqrt{t} - \sqrt{t}}^A \right] \]

where \( Y_{b\sqrt{t} - \sqrt{t}}^A \) is the truncated McKean martingale defined in (9.2.29). Note that its second moment is bounded by \( D_2(r) \) (see (9.2.43)). Comparing this to (9.2.73), one sees that

\[ \frac{\mathbb{E}[x^2 | \tilde{h}_{b\sqrt{t}}]}{\mathbb{E}[x | \tilde{h}_{b\sqrt{t}}]} \leq D_2(r) e^{-(1 + \sigma^2_t) b\sqrt{t} + \sqrt{t}z_t(b\sqrt{t})} \leq Ce^{-(1 - \sigma^2_t) b\sqrt{t} + o(1)}, \quad (9.2.74) \]

which tends to zero uniformly as \( t \uparrow \infty \). Thus the second moment term is negligible. Hence we only have to control

\[ \prod_{1 \leq i \leq n(b\sqrt{t})} \left( 1 - C(a) e^{-(1 + \sigma^2_t) b\sqrt{t} + \sqrt{t}z_t(b\sqrt{t}) - \sqrt{t}y} \sigma_2 \right) \]

\[ = \mathbb{E} \left[ \exp \left( - \sum_{1 \leq i \leq n(b\sqrt{t})} C(a) e^{-(1 + \sigma^2_t) b\sqrt{t} + \sqrt{t}z_t(b\sqrt{t}) - \sqrt{t}y} \sigma_2 \right) (1 + o(1)) \right] \]

\[ = \mathbb{E} \left[ \exp \left( -C(a) \sigma_2 e^{-\sqrt{t}Y_{b\sqrt{t},y}^B} \right) (1 + o(1)) \right] \quad (9.2.75) \]

where
\[
\tilde{Y}^B_{b\sqrt{t},Y} = \sum_{i=1}^{n(b\sqrt{t})} e^{-(1+\sigma_1^2)b\sqrt{t}+\sqrt{\zeta_n(b\sqrt{t})} |_{x_i-(b\sqrt{t})-\sqrt{\zeta_n^2 b\sqrt{t}}\in[-Bt^{1/2},Bt^{1/2}]}}, \tag{9.2.76}
\]

Now from Lemma 9.11, \(\tilde{Y}^B_{b\sqrt{t},Y}\) converges in probability and in \(L^1\) to the random variable \(Y\), when we let first \(t\) and then \(B\) tend to infinity. Since \(\tilde{Y}^B_{b\sqrt{t},Y} \geq 0\) and \(C(a) > 0\), it follows

\[
\lim_{B\to\infty} \liminf_{t\to\infty} \mathbb{E} \left[ \exp \left( -C(a) \sigma_2 \tilde{Y}^B_{b\sqrt{t},Y} e^{-\sqrt{\zeta}} \right) \right] = \mathbb{E} \left[ \exp \left( -\sigma_2 C(a) Y e^{-\sqrt{\zeta}} \right) \right]. \tag{9.2.77}
\]

Finally, letting \(r\) tend to \(+\infty\), all the \(\varepsilon\)-errors (that are still present implicitly, vanish. This concludes the proof of Theorem 9.15. \(\Box\)

### 9.2.5 Existence of the limiting process

The following existence theorem is the basic step in the proof of Theorem 9.2.1.

**Theorem 9.16.** Let \(\sigma_1 < \sigma_2\). Then, the point processes \(\delta_t = \sum_{k \leq n(t)} \delta_{x(k(t)-m(t))}\) converges in law to a non-trivial point process \(\delta\).

It suffices to show that, for \(\phi \in C_c(\mathbb{R})\) positive, the Laplace functional

\[
\Psi_t(\phi) = \mathbb{E} \left[ \exp \left( -\int \phi(y) \delta_t(dy) \right) \right], \tag{9.2.78}
\]

of the processes \(\delta_t\) converges. The proof of this is essentially a combination of the corresponding proof for ordinary BBM and what we did when showing convergence of the maximum. The result is that

\[
\lim_{t\to\infty} \Psi_t(\phi) = \mathbb{E} \left[ \exp \left( -\sigma_2 C(a, \phi) Y \right) \right], \tag{9.2.79}
\]

where \(C(a, \phi)\) is given by

\[
C(a, \phi) = \lim_{t\to\infty} \frac{1}{\sqrt{2\pi}} \int_0^\infty v(t, z + \sqrt{2}t)e^{(\sqrt{2}+\phi)z-a^2t/2} dz, \tag{9.2.80}
\]

where \(v(t, x)\) is the solutions of the F-KPP equation with initial data \(v(0, x) = 1 - e^{-\phi(-x)}\). Here \(\phi(z) \equiv \phi(\sigma_2 z)\).

### 9.2.6 The auxiliary process

The final step is the interpretation of the limiting process. This is again very much analogous to the standard BBM case. Let \((\eta_i; i \in \mathbb{N})\) be the atoms of a Poisson point
process \( \eta \) on \((-\infty, 0)\) with intensity measure
\[
\frac{\sigma^2}{\sqrt{2\pi}} e^{-(\sqrt{2}+a)z} e^{-a^2/2} dz.
\]
(9.2.81)

For each \( i \in \mathbb{N} \) consider independent standard BBMs \( \bar{x}^i \). The auxiliary point process of interest is the superposition of the i.i.d BBMs with drift shifted by \( \eta_i + \frac{1}{\sqrt{2}+a} \log Y \), where \( a \) is the constant defined in (9.2.60):
\[
\Pi_t = \sum_{i,k} \delta_{\eta_i + \frac{1}{\sqrt{2}+a} (\log Y + \bar{x}^i_k(t)) - \sqrt{2}} \sigma^2.
\]
(9.2.82)

**Remark 9.17.** The form of the auxiliary process is similar to the case of standard BBM, but with a different intensity of the Poisson process. In particular, the intensity decays exponentially with \( t \). This is a consequence of the fact that particles at the time of the speed change were forced to be \( O(t) \) below the line \( \sqrt{2}t \), in contrast to the \( O(\sqrt{t}) \) in the case of ordinary BBM. The reduction of the intensity of the process with \( t \) forces the particles to be selected at these locations.

**Theorem 9.18.** Let \( \delta_t \) be the extremal process of the two-speed BBM. Then
\[
\lim_{t \to \infty} \delta_t \overset{law}{=} \lim_{t \to \infty} \Pi_t.
\]
(9.2.83)

As the proof is in nature similar to that in the standard case, we skip the details.

The following proposition shows that in spite of the different Poisson ingredients, when we look at the process of the extremes of each of the \( x^i(t) \), we end up with a Poisson point process just like in the standard BBM case.

**Proposition 9.19.** Define the point process
\[
\Pi_t^{ext} \equiv \sum_{i,k} \delta_{\eta_i + \frac{1}{\sqrt{2}+a} \log Y + \max_{k \leq p(i)} \bar{x}^i_k(t) - \sqrt{2}} \sigma^2.
\]
(9.2.84)

Then
\[
\lim_{t \to \infty} \Pi_t^{ext} \overset{law}{=} P_Y \equiv \sum_{i \in \mathbb{N}} \delta_{p(i)},
\]
(9.2.85)

where \( P_Y \) is the Poisson point process on \( \mathbb{R} \) with intensity measure \( \sigma^2 C(a) Y \sqrt{2} e^{-\sqrt{2}x} dx \).

**Proof.** We consider the Laplace functional of \( \Pi_t^{ext} \). Let \( M^{(i)}(t) = \max_{k \leq p(i)} x_k^{(i)}(t) \) and as before \( \tilde{\phi}(z) = \phi(\sigma^2 z) \). We want to show
\[
\lim_{t \to \infty} \mathbb{E} \left[ \exp \left( - \sum_i \tilde{\phi}(\eta_i + M^{(i)}(t) - \sqrt{2}t) \right) \right]
= \exp \left( -\sigma^2 C(a) \int_{-\infty}^{\infty} \left( 1 - e^{-\phi(s)} \right) \sqrt{2} e^{-\sqrt{2}x} dx \right).
\]
(9.2.86)
Since \( \eta \) is a Poisson point process and the \( M^{(i)} \) are i.i.d. we have

\[
\mathbb{E} \left[ \exp \left( -\sum_i \tilde{\phi}(\eta_i + M^{(i)}(t) - \sqrt{2} t) \right) \right]
\]

\[
= \exp \left( -\sigma^2 \int_0^\infty \mathbb{E} \left[ 1 - e^{-\tilde{\phi}(z + M(t) - \sqrt{2} z)} \right] e^{-\left(\sqrt{2} + a\right) z - a t^2 / 2} \frac{dz}{\sqrt{2\pi}} \right), \tag{9.2.87}
\]

where \( M(t) \) has the same distribution as one the variables \( M^{(i)}(t) \). Now we apply Lemma 9.14 with \( h(x) = 1 - e^{-\tilde{\phi}(x)} \). Hence the result follows by using that \( \tilde{\phi}(z) = \phi(\sigma z) \) and \( \sqrt{2} + a = \sqrt{2} \sigma \) together with the change of variables \( x = \sigma z \). \( \Box \)

The following proposition states that the Poisson points of the auxiliary process contribute to the limiting process come from a neighbourhood of \(-at\).

**Proposition 9.20.** Let \( z \in \mathbb{R}, \varepsilon > 0 \). Let \( \eta_i \) be the atoms of a Poisson point process with intensity measure \( C e^{-\left(\sqrt{2} + a\right) x - a t^2 / 2} \) on \((\infty, 0] \). Then there exists \( B < \infty \) such that

\[
\sup_{t \geq 0} \mathbb{P} \left( \exists i, k : \eta_i + x_i^{(i)}(t) - \sqrt{2} t \geq z, \eta_i \not\in \left[-at - B\sqrt{t}, -at + B\sqrt{t}\right] \right) \leq \varepsilon. \tag{9.2.88}
\]

**Proof.** By a first order Chebychev inequality we have

\[
\mathbb{P} \left( \exists i, k : \eta_i + x_i^{(i)}(t) - \sqrt{2} t \geq z, \eta_i > -at + B\sqrt{t} \right)
\]

\[
\leq C \int_{-at + B\sqrt{t}}^0 \mathbb{P} \left( \max_i x_i(t) \geq \sqrt{2} t + x + z \right) e^{-\left(\sqrt{2} + a\right) x - a t^2 / 2} dx
\]

\[
= C \int_{-at}^{at - B\sqrt{t}} \mathbb{P} \left( \max_i x_i(t) \geq \sqrt{2} t + x + z \right) e^{\left(\sqrt{2} + a\right) x} e^{-a t^2 / 2} dx, \tag{9.2.89}
\]

by the change of variables \( x \rightarrow -x \). Using the asymptotics of Lemma 6.45 we can bound (9.2.89) from above by

\[
\rho Ce^{-\left(\sqrt{2} + a\right) z} \int_{-a\sqrt{t}}^{at - B\sqrt{t}} e^{-x^2 / 2} dx \leq \rho Ce^{-\left(\sqrt{2} + a\right) z} e^{-B^2 / 2}, \tag{9.2.90}
\]

by changing variables \( x \rightarrow x / \sqrt{t} - a \sqrt{t} \). For any \( z \in \mathbb{R} \), this can be made as small as we want by choosing \( B \) large enough. Similarly one bounds

\[
\mathbb{P} \left( \exists i, k : \eta_i + x_i^{(i)}(t) - \sqrt{2} t \geq z, \eta_i < -at - B\sqrt{t} \right) \leq \rho Ce^{-\left(\sqrt{2} + a\right) z} e^{-B^2 / 2}. \tag{9.2.91}
\]

This concludes the proof. \( \Box \)

The next proposition describes the law of the clusters \( \bar{x}_k^{(i)} \). This is analogous to Theorem 7.16 (Theorem 3.4 in [6]).
**Proposition 9.21.** Let \( x = at + o(t) \) and \( \{ \bar{x}_k(t), k \leq n(t) \} \) be a standard BBM under the conditional law \( \mathbb{P} \left( \cdot \mid \{ \max \bar{x}_k(t) - \sqrt{2}t - x > 0 \} \right) \). Then the point process

\[
\sum_{k \leq n(t)} \delta_{\bar{x}_k(t) - \sqrt{2}t - x} \tag{9.2.92}
\]

converges in law under \( \mathbb{P} \left( \cdot \mid \{ \max \bar{x}_k(t) - \sqrt{2}t - x > 0 \} \right) \) as \( t \to \infty \) to a well defined point process \( \bar{\xi} \). The limit does not depend on \( x \) and the maximum of \( \bar{\xi} \) shifted by \( x \) has the law of an exponential random variable with parameter \( \sqrt{2} + a \).

**Proof.** Set \( \bar{\xi} = \sum_k \delta_{\bar{x}_k(t) - \sqrt{2}t} \) and \( \max \bar{\xi} = \max \bar{x}_k(t) - \sqrt{2}t \). First we show that for \( X > 0 \)

\[
\lim_{t \to \infty} \mathbb{P} \left( \max \bar{\xi} > X + x \mid \max \bar{\xi} > x \right) = e^{-\sqrt{2} + a}X, \tag{9.2.93}
\]

To see this we rewrite the conditional probability as \( \frac{\mathbb{P}[\max \bar{\xi} > X + x]}{\mathbb{P}[\max \bar{\xi} > x]} \) and use the uniform bounds from Proposition 6.2. Observing that

\[
\lim_{t \to \infty} \frac{\Psi(r,t,X + x + \sqrt{2}t)}{\Psi(r,t,x + \sqrt{2}t)} = e^{-\sqrt{2} + a}X, \tag{9.2.94}
\]

where \( \Psi \) is defined in Equation (9.2), we get (9.2.93) by first taking \( t \to \infty \) and then \( r \to \infty \). The general claim of Proposition 9.21 follows in exactly the same way from (9.2.93) as Theorem 7.16. \( \square \)

Define the gap process

\[
D_t = \sum_k \delta_{\bar{x}_k(t) - \max_j \bar{x}_j(t)}. \tag{9.2.95}
\]

Denote by \( \bar{\xi}_i \) the atoms of the limiting process \( \bar{\xi} \), i.e. \( \bar{\xi} = \sum_j \delta_{\bar{\xi}_j} \) and define

\[
D = \sum_j \delta_{\Lambda_j}, \quad \Lambda_j = \bar{\xi}_j - \max_{i} \bar{\xi}_i. \tag{9.2.96}
\]

\( D \) is a point process on \((-\infty, 0]\) with an atom at 0.

**Corollary 9.22.** Let \( x = at + o(t) \). In the limit \( t \to \infty \) the random variables \( D_t \) and \( x + \max \bar{\xi} \) are conditionally independent on the event \( \{ x + \max \bar{\xi} > b \} \) for any \( b \in \mathbb{R} \). More precisely, for any bounded function \( f, h \) and \( \bar{\phi} \in \mathcal{C}_r(\mathbb{R}) \),

\[
\lim_{t \to \infty} \mathbb{E} \left[ f \left( \int \bar{\phi}(z) D_t(dz) \right) h(x + \max \bar{\xi}) \mid x + \max \bar{\xi} > b \right] = \mathbb{E} \left[ f \left( \int \bar{\phi}(z) D(dz) \right) \right] \frac{\int_0^b h(z)(\sqrt{2} + a)e^{-(\sqrt{2} + a)z}dz}{e^{-(\sqrt{2} + a)b}}. \tag{9.2.97}
\]

**Proof.** The proof is essentially identical to the proof of Corollary 4.12 in [6]. \( \square \)
Finally we come to the description of the extremal process as seen from the Poisson process of cluster extremes, which is the formulation of Theorem 9.2.1

**Theorem 9.23.** Let \( P_Y \) be as in (9.2.85) and let \( \{ D^{(i)}, i \in \mathbb{N} \} \) be a family of independent copies of the gap-process (9.2.96) with atoms \( \Lambda^{(i)}_j \). Then the point process \( \mathcal{E} \) converges in law as \( t \to \infty \) to a Poisson cluster point process \( \mathcal{E} \) given by

\[
\mathcal{E} \overset{law}{=} \sum_{i,j} \delta_{p_i+\sigma_2 \Lambda^{(i)}_j}.
\]

Also this proof is now very close to that of Theorem 7.15, resp. Theorem 2.1 in [6].

### 9.2.7 The case \( \sigma_1 > \sigma_2 \)

In this section we proof Theorem 9.2. The existence of the process \( \mathcal{E} \) from (9.2.11) will be a byproduct of the proof.

The point is that in this case if we rerun the computations to find the location of the particles that contribute to the maximum at time \( bt \), the naive computation would place them above the level of the maximal particles at that time. But of course, there are no such particles. Thus the particles that reach the highest level are those that have been maximal at time \( bt \). The following lemma that is contained in the calculation of the maximal displacement in [35] makes this precise.

**Lemma 9.24.** [35] For all \( \varepsilon > 0, d \in \mathbb{R} \) there exists a constant \( D \) large enough such that for \( t \) sufficiently large

\[
P[\exists k \leq n(t) : x_k(t) > m(t) + d \text{ and } x_k(bt) < m_1(bt) - D] < \varepsilon.
\]

**Proof (of Theorem 9.2).** First we establish the existence of a limiting process. Note that \( m(t) = m_1(bt) + m_2((1-b)t) \), where \( m_i(s) = \sqrt{2} \sigma_i s - \frac{1}{2 \sqrt{2}} \sigma_i \log s \). Recall

\[
\tilde{\phi}(z) = \phi(\sigma_2 z)
\]

and

\[
g(z) = 1 - e^{-\tilde{\phi}(-z)}.
\]

Here we assume that \( \phi(x) = 1_{x > a} \) for \( a \in \mathbb{R} \). Using that the maximal displacement is \( m(t) \) in this case we can proceed as in the proof of Theorem 9.16 and only have to control

\[
1 - \Psi(t) = \mathbb{E} \left[ \prod_{j \leq m(t)} \mathbb{E} \left[ \prod_{j \leq n_j((1-b)t)} g((m(t) - x_j(bt))/\sigma_2 - x'_j((1-b)t)) \right] \right],
\]

\[
(9.2.102)
\]
where $\tilde{x}_j((1-b)t)$ are the particles of a standard BBM at time $(1-b)t$ and $x_j(bt)$ are the particles of a BBM with variance $\sigma_j$ at time $bt$. Using Lemma 9.24 and Theorem 1.2 of [35] as in the proof of Theorem 9.15 above, we obtain that (9.2.102), for $t$ sufficiently large, equals

$$E \left[ \prod_{i \leq n((1-b)t)} \prod_{x_i(bt) > m_1(bt) - D} g \left( \frac{(m(t)-x_i(bt))}{\sigma_2} - \tilde{x}_j((1-b)t) \right) \bigg| \bar{F}_{bt} \right] + O(\epsilon).$$

(9.2.103)

The remainder of the proof has an iterated structure. In a first step we show that conditioned on $\bar{F}_{bt}$ for each $i \leq n(bt)$ the points $\{x_i(bt) + x_j((1-b)t) - m(t)|x_i(bt) > m_1(bt) - D\}$ converge to the corresponding points of the point process $x_i(bt) - m_1(bt) + \sigma_2 \tilde{\varepsilon}^{(i)}$, where $\tilde{\varepsilon}^{(i)}$ are independent copies of the extremal process of standard BBM. To this end observe that

$$u((1-b)t,z) = E \left[ \prod_{i \leq n((1-b)t)} g(z - \tilde{x}_j((1-b)t)) \right]$$

(9.2.104)

solves the F-KPP equation with initial condition $u(0,z) = g(z)$. Moreover, the assumptions of Theorem ?? are satisfied. Hence (9.2.103) is equal to

$$\epsilon + E \left[ \prod_{i \leq n(bt)} \prod_{x_i(bt) > m_1(bt) - D} \left( E \left[ e^{C(\bar{\phi})Ze^{-\sqrt{\frac{m_1(bt)-x_i(bt)}{\sigma_2}}} | \bar{F}_{bt} \right] (1 + o(1)) \right) \right].$$

(9.2.105)

Here $C(\bar{\phi})$ is from standard BBM, i.e.

$$C(\bar{\phi}) = \lim_{t \uparrow \infty} \sqrt{\frac{2}{\pi}} \int_0^\infty u(t,y + \sqrt{2t})ye^{\sqrt{2t}y}dy,$$

(9.2.106)

see Lemma ??, Note furthermore that already in (9.2.105) the concatenated structure of the limiting point process becomes visible. In a second step we establish that the points $x_i(bt) - m_1(t)$ that have a descendant in the lead at time $t$ converge to $\tilde{\varepsilon}$.

Define

$$h_D(y) = \begin{cases} E \left[ \exp \left( -C(\bar{\phi})Ze^{-\sqrt{\frac{m_1(bt)}{\sigma_2}}} \right) \right], & \text{if } \sigma_1 y < D, \\ 1, & \text{if } \sigma_1 y \geq D. \end{cases}$$

(9.2.107)

Then the expectation in (9.2.105) can be written as (we ignore the error term $o(1)$ which is easily controlled using that the probability that the number of terms in the product is larger than $N$ tends to zero as $N \uparrow \infty$, uniformly in $t$)
Two-speed BBM

\[ E \left[ \prod_{i \leq n(bt)} h_{\delta,D}(m_1(bt) \sigma_1 - \bar{x}_i(t)) \right], \quad (9.2.108) \]

where now \( \bar{x} \) is standard BBM. Defining

\[ v_D(t,z) = 1 - E \left[ \prod_{i \leq n(t)} h_D(z - \bar{x}_i(bt)) \right], \quad (9.2.109) \]

\( v_D \) is a solution of the F-KPP equation with initial condition \( v_D(0,z) = 1 - h_D(z) \). But this initial condition satisfies the assumptions of Theorem ??? and therefore,

\[ v_D(t,m(t) + x) \to E \left[ e^{-\tilde{C}(D,Z,C(\tilde{\phi}))2\sigma_1} \right]. \quad (9.2.110) \]

where \( \tilde{Z} \) is an independent copy of \( Z \) and

\[ \tilde{C}(D,Z,C(\tilde{\phi})) = \lim_{D \to \infty} \sqrt{\frac{2}{\pi}} \int_0^\infty v_D(t,y + \sqrt{2t})ye^{\gamma Z_t}dy. \quad (9.2.111) \]

By the same argumentation as in standard BBM setting one obtains that

\[ \tilde{C}(Z,C(\tilde{\phi})) = \lim_{D \to \infty} \tilde{C}(D,Z,C(\tilde{\phi})) = \lim_{t \to \infty} \sqrt{\frac{2}{\pi}} \int_0^\infty v(t,y + \sqrt{2t})ye^{\gamma Z_t}dy, \quad (9.2.112) \]

where \( v \) is the solution of the F-KPP equation with initial condition \( v(0,z) = 1 - h(z) \) with

\[ h(z) = E \left[ e^{-C(\tilde{\phi})Z e^{-\gamma Z_t}} \right]. \quad (9.2.113) \]

Therefore, taking the limit first as \( D \to \infty \) in the left-hand side of (9.2.110), we get that

\[ \lim_{t \to \infty} \Psi_t(\phi(\cdot + x)) = \lim_{D \to \infty} \lim_{t \to \infty} v_D(t,m(t) + x) = E \left[ e^{-\tilde{C}(Z,C(\tilde{\phi}))2e^{-\gamma Z_t}} \right]. \quad (9.2.114) \]

To see that the constants \( \tilde{C}(Z,C(\tilde{\phi})) \) are strictly positive, one uses the Laplace functionals \( \Psi(t) \) are bounded from above by

\[ E \left[ \exp \left( -\phi \left( \max_{i \leq n(bt)} x_i(bt) + \max_{j \leq n_1((1-b)t)} x_j^1((1-b)t) - m(t) \right) \right) \right]. \quad (9.2.115) \]

Here we used that the offspring of any of the particles at time \( bt \) has the same law. So the sum of the two maxima in the expression above has the same distribution as the largest descendent at time \( t \) off the largest particle at time \( bt \). The limit of Eq. (9.2.115) as \( t \to \infty \) exists and is strictly smaller than 1 by the convergence in law of the recentered maximum of a standard BBM. But this implies the positivity of the constants \( \tilde{C} \). Hence a limiting point process exists. Finally, one may easily check that the right hand side of (9.2.114) coincides with the Laplace functional of
the point process defined in (9.2.11) by basically repeating the computations above.

\[ \square \]

**Remark 9.25.** Note that in particular, the structure of the variance profile is contained in the constant \( \tilde{C}(D, Z, C(\tilde{\phi})) \) and that also the information on the structure of the limiting point process is contained in this constant. In fact, we see that in all cases we have considered in this paper, the Laplace functional of the limiting process has the form

\[
\lim_{t \to \infty} \Psi_t(\phi(\cdot + x)) = E \left[ \exp \left( -C(\phi) M e^{-\sqrt{2}x} \right) \right], \tag{9.2.116}
\]

where \( M \) is a martingale limit (either \( Y \) of \( Z \)) and \( C \) is a map from the space of positive continuous functions with compact support to the real numbers. This function contains all the information on the specific limiting process. This is compatible with the finding in [59] in the case where the speed is a concave function of \( s/t \). The universal form (9.2.116) is thus misleading and without knowledge of the specific form of \( C(\phi) \), (9.2.116) contains almost no information.

**Remark 9.26.** There is no difficulty to extend this result to the case of multi-speed BBM with decreasing speeds.

### 9.3 Universality below the straight line

It turns out that in the case when the covariance stays strictly below the straight line \( A(s) = s \) for all \( s \in (0, t) \), the same picture emerges as in the two speed case, with the McKean martingale depending only on the slope at 0 and the decoration process depending only on the slope at 1.

In [18] a rather large class of functions \( A \) were considered. Here we will simplify the presentation by restricting ourselves to smooth functions. Let \( A : [0, 1] \to [0, 1] \) be non-decreasing function twice differentiable with bounded second derivative that satisfies the following three conditions:

\begin{enumerate}
\item[(A1)] For all \( x \in (0, 1) \): \( A(x) < x \), \( A(0) = 0 \) and \( A(1) = 1 \).
\item[(A2)] \( A'(0) = \sigma_b^2 < 1 \) and \( A'(1) = \sigma_e^2 > 1 \).
\end{enumerate}

**Theorem 9.27.** Assume that \( A : [0, 1] \to [0, 1] \) satisfies (A1)-(A2). Let \( \tilde{m}(t) = \sqrt{2t} - \frac{1}{2\sqrt{2}} \log t \). Then there is a constant \( \tilde{C}(\sigma_e) \) depending only on \( \sigma_e \) and a random variable \( Y_{\sigma_b} \) depending only on \( \sigma_b \) such that

\[
(i) \quad \lim_{t \to \infty} \mathbb{P} \left( \max_{1 \leq i \leq n(t)} x_i(t) - \tilde{m}(t) \leq x \right) = E \left[ e^{-\tilde{C}(\sigma_e) Y_{\sigma_b} e^{-\sqrt{2}t}} \right], \tag{9.3.1}
\]

\[
(ii) \quad \text{The point process}
\]
9.3 Universality below the straight line

\[ \sum_{\delta_{\tilde{n}}(t) - \tilde{v}(t)} \delta_{\sigma_b, \sigma_e} = \sum_{i,j} \delta_{p_i + \sigma_e \Lambda^{(i)}}; \quad (9.3.2) \]

as \( t \uparrow \infty \), in law, where the \( p_i \) are the atoms of a Poisson point process on \( \mathbb{R} \) with intensity measure \( C(\sigma_e)Y_{\sigma_b}e^{-\sqrt{2\pi}t}dx \), and the \( \Lambda^{(i)} \) are the limits of the processes as in (\ref{eq:Lambda-limit})

(iii) If \( \Lambda'(1) = \infty \), then \( \tilde{C}(\infty) = 1/\sqrt{4\pi} \), and \( \Lambda^{(i)} = \delta_0 \), i.e. the limiting process is a Cox process.

The random variable \( Y_{\sigma_b} \) is the limit of the uniformly integrable martingale

\[ Y_{\sigma_b}(s) = \sum_{i=1}^{\pi(s)} e^{-s(1+\sigma_i^2)+\sqrt{2\sigma_b}\bar{\xi}_i(s)}, \quad (9.3.3) \]

where \( \bar{\xi}_i(s) \) is standard branching Brownian motion.

9.3.1 Outline of the proof

The proof of Theorem 9.27 is based on the corresponding result obtained in \cite{17} for the case of two speeds, and on a Gaussian comparison method. We start by showing the localisation of paths, namely that the paths of all particles that reach a height of order \( \tilde{m}(t) \) at time \( t \) has to lie within a certain tube. Then we show tightness of the extremal process.

The remainder of the work consists in proving the convergence of the finite dimensional distributions. To this end we use Laplace transforms. We introduce auxiliary two speed BBM's whose covariance functions approximate \( \Sigma^2(s) \) well around 0 and \( t \). Moreover we choose them in such a way that their covariance functions lie above respectively below \( \Sigma^2(s) \) in a neighbourhood of 0 and \( t \).

We then use Gaussian comparison methods to compare the Laplace transforms. The Gaussian comparisons comes in three main steps. In a first step we introduce the usual interpolating process and introduce a localisation condition on its paths. In a second step we justify a certain integration by parts formula, that is adapted to our setting. Finally the resulting quantities are decomposed in a part that has a sign and a part that converges to zero.
9.3.2 Localization of paths

An important first step is again the localisation of the ancestral paths of particles that reach extreme levels. This is essentially inherited from properties of the standard Brownian bridge. For a given covariance function $\Sigma^2$, and a subinterval $I \subset [0,t]$, define the following events on the space of paths, $X : \mathbb{R}_+ \to \mathbb{R}$,

$$\mathcal{G}^\gamma_{I,s} = \left\{ X \mid \forall s \in I : \left| X(s) - \frac{\Sigma^2(s)}{t} X(t) \right| < (\Sigma^2(s) \wedge (t - \Sigma^2(s)))^\gamma \right\}.$$  

(9.3.4)

**Proposition 9.28.** Let $x$ denote the variable speed BBM with covariance function $\Sigma^2$. For any $\frac{1}{2} < \gamma < 1$ and for all $d \in \mathbb{R}$, there exists $r$ sufficiently large such that, for all $t > 3r$,

$$\mathbb{P} \left( \exists k \leq n(t) : \{ x_k(t) > \tilde{m}(t) + d \} \wedge \{ x_k \notin \mathcal{G}^\gamma_{I,s} \} \right) < \varepsilon,$$

(9.3.5)

where $I_r \equiv \{ s : \Sigma^2(s) \in [r,t-r] \}$.

To prove Proposition 9.28 we need Lemma 6.29 on Brownian bridges.

**Proof (Proof of Proposition 9.28).** Using a first moment method, the probability in (9.3.5) is bounded from above by

$$e^t \mathbb{P} \left( B_{\Sigma^2(t)} > \tilde{m}(t) + d, B_{\Sigma^2(t)} \notin \mathcal{G}^\gamma_{I,s} \right),$$

(9.3.6)

where $B_{\Sigma^2(t)}$ is a time change of an ordinary Brownian motion. Using that $\Sigma^2(s)$ is an non-decreasing function on $[0,t]$ with $\Sigma^2(t) = t$, we bound (9.3.6) from above by

$$e^t \mathbb{P} \left( \{ B_t > \tilde{m}(t) + d \} \wedge \{ \exists s \in [r,t-r] : \left| B_s - \frac{s}{t} B_t \right| > (s \wedge (t-s))^\gamma \} \right).$$

(9.3.7)

Now, $\xi(s) \equiv B_s - \frac{s}{t} B_t$ is a Brownian bridge from 0 to 0 in time $t$, and it is well known that $\xi(s)$ is independent of $B_t$. Therefore, it holds that (9.3.7) is equal to

$$e^t \mathbb{P}(B_t > \tilde{m}(t) + d) \mathbb{P}(\exists s \in [r,t-r] : |\xi(s)| > (s \wedge (t-s))^\gamma).$$

(9.3.8)

Using the standard Gaussian tail bound,

$$\int_u^\infty e^{-x^2/2} dx \leq u^{-1} e^{-u^2/2}, \quad \text{for } u > 0,$$

(9.3.9)

we have

$$e^t \mathbb{P}(B_t > \tilde{m}(t) + d) \leq e^{t} \frac{\sqrt{t}}{\sqrt{2\pi (\tilde{m}(t) + d)}} e^{-\tilde{m}(t) + d^2/2t}$$

$$= \frac{1}{\sqrt{2\pi (\tilde{m}(t) + d)}} e^{-\sqrt{2d}} \leq M,$$

(9.3.10)
for some constant $M > 0$. By Lemma 6.29 we can find by $r$ large enough such that
\[ P \left( \exists s \in [r, t - r] : |\xi(s)| > (s \wedge (t - s))^\gamma \right) < \varepsilon / M. \] (9.3.11)

Using the bounds of (9.3.10) and (9.3.10) we can bound (9.3.8) from above by $\varepsilon$.

**Proof (Proof of Theorem 9.27).** We show the convergence of the extremal process
\[
E_t = \sum_{k \leq n(t)} \delta_{x_k(t) - \tilde{m}(t)},
\] (9.3.12)
by showing the convergence of the finite dimensional distributions and tightness. Tightness of $(E_t)_{t \geq 0}$ follows trivially from a first order Chebyshev estimate that shows that for any $d \in \mathbb{R}$ and $\varepsilon > 0$, there exists $N = N(\varepsilon, d)$ such that, for all $t > 0$,
\[ P(E_t[d, \infty) \geq N) < \varepsilon. \] (9.3.13)

To show the convergence of the finite dimensional distributions define, for $u \in \mathbb{R}$,
\[
N_u(t) = \sum_{i=1}^{n(t)} \mathbb{1}_{x_i(t) - \tilde{m}(t) > u},
\] (9.3.14)
that counts the number of points that lie above $u$. Moreover, we define the corresponding quantity for the process $E_{\sigma_b, \sigma_e}$ (defined in (9.3.2)),
\[
N'_u = \sum_{i,j} \mathbb{1}_{p_i + \sigma_e \Lambda(i) > u}.
\] (9.3.15)

Observe that, in particular,
\[ P \left( \max_{1 \leq k \leq n(t)} x_k(t) - \tilde{m}(t) \leq u \right) = P \left( N_u(t) = 0 \right). \] (9.3.16)

The key step in the proof of Theorem 9.27 is the following proposition, that asserts the convergence of the finite dimensional distributions of the process $E_t$.

**Proposition 9.29.** For all $k \in \mathbb{N}$ and $u_1, \ldots, u_k \in \mathbb{R}$
\[ \{N_{u_1}(t), \ldots, N_{u_k}(t)\} \overset{d}{\to} \{N_{u_1}, \ldots, N_{u_k}\} \] (9.3.17)
as $t \uparrow \infty$.

The proof of this proposition will be postponed to the following sections.
Assuming the proposition, we can now conclude the proof of the theorem.
(9.3.17) implies the convergence of the finite dimensional distributions of $E_t$.
Since tightness follows from the control of the first moment, we obtain Assertion (ii) of Theorem 9.27. Assertion (i) follows immediately from Eq. (9.3.16).
To prove Assertion (iii), we need to show that, as \( \sigma_e^2 \uparrow \infty \), it holds that \( \tilde{C}(\sigma_e) \uparrow 1/\sqrt{4\pi} \) and the processes \( A^{(i)} \) converge to the trivial process \( \delta_0 \). Then,
\[
e_{\sigma_e,\infty} = \sum_i \delta_{p_i},
\]
where \( (p_i, i \in \mathbb{N}) \) are the points of a PPP with random intensity measure \( \frac{1}{\sqrt{4\pi}} y_0 e^{-\sqrt{2}y} dy \).

**Lemma 9.30.** The point process \( \mathcal{E}_{\sigma_e,\sigma_e} \) converges in law, as \( \sigma_e \uparrow \infty \), to the point process \( \mathcal{E}_{\sigma_e,\infty} \).

**Proof.** The proof of Lemma 9.30 is based on a result concerning the cluster processes \( A^{(i)} \). We write \( \Lambda_{\sigma_e} \) for a single copy of these processes and add the subscript to make the dependence on the parameter \( \sigma_e \) explicit. We recall from [17] that the process \( \Lambda_{\sigma_e} \) is constructed as follows. Define the processes \( \mathcal{E}_{\sigma_e} \) as the limits of the point processes
\[
\mathcal{E}_{\sigma_e} = \sum_{k=1}^{n(t)} \delta_{x_k(t) - \sqrt{2}\sigma_e t}
\]
where \( x \) is standard BBM at time \( t \) conditioned on the event \( \{ \max_{k \leq n(t)} x_k(t) > \sqrt{2}\sigma_e t \} \). We show here that, as \( \sigma_e \) tends to infinity, the processes \( \mathcal{E}_{\sigma_e} \) converge to a point process consisting of a single atom at 0. More precisely, we show that
\[
\lim_{\sigma_e \uparrow \infty} \lim_{t \to \infty} \mathbb{P} \left( \mathcal{E}_{\sigma_e}([-R, \infty)) > 1 \mid \max_{k \leq n(t)} x_k(t) > \sqrt{2}\sigma_e t \right) = 0.
\]

Now,
\[
\mathbb{P} \left( \mathcal{E}_{\sigma_e}([-R, \infty)) > 1 \mid \max_{k \leq n(t)} x_k(t) > \sqrt{2}\sigma_e t \right)
\]
\[
\leq \mathbb{P} \left( \text{supp} \mathcal{E}_{\sigma_e} \cap [0, \infty) \neq \emptyset \land \mathcal{E}_{\sigma_e}([-R, \infty)) > 1 \mid \max_{k \leq n(t)} x_k(t) > \sqrt{2}\sigma_e t \right)
\]
\[
\leq \int_0^\infty \mathbb{P} \left( \text{supp} \mathcal{E}_{\sigma_e} \cap dy \neq \emptyset \land \mathcal{E}_{\sigma_e}([-R, \infty)) > 1 \mid \max_{k \leq n(t)} x_k(t) > \sqrt{2}\sigma_e t \right) dy
\]
\[
= \int_0^\infty \mathbb{P} \left( \text{supp} \mathcal{E}_{\sigma_e} \cap dy \neq \emptyset \right) \mathbb{P} \left( \max_{k \leq n(t)} x_k(t) > \sqrt{2}\sigma_e t \right) dy \times \mathbb{P} \left( \mathcal{E}_{\sigma_e}([-R, \infty)) > 1 \mid \text{supp} \mathcal{E}_{\sigma_e} \cap dy \neq \emptyset \right).
\]

But \( \mathbb{P} \left( \text{supp} \mathcal{E}_{\sigma_e} \cap dy \neq \emptyset \right) \equiv P_{t,y+\sqrt{2}\sigma_e} \) is the Palm measure on BBM, i.e. the conditional law of BBM given that there is a particle at time \( t \) in \( dy \) (see Kallenberg [50] Theorem 12.8]. Chauvin, Rouault, and Wakolbinger [29] Theorem 2] describe the tree under the Palm measure \( P_{t,z} \) as follows. Pick one particle at time \( t \) at the location \( z \). Then pick a spine \( Y \), which is a Brownian bridge from 0 to \( z \) in time \( t \). Next pick a Poisson point process \( \pi \) on \([0,t]\) with intensity 2. For each point \( p \in \pi \)
start a random number $\nu_p$ of independent branching Brownian motions $(B^{(p),i}Y, i \leq \nu_p)$ starting at $Y(p)$. The law of $\nu$ is given by the size biased distribution, $\mathbb{P}(\nu_p = k - 1) \sim \frac{k_p}{\nu_p}$. See Figure 9.1. Now let $z = \sqrt{2} \sigma_t + y$ for $y \geq 0$. Under the Palm measure, the point process $E_{\sigma_t}(t)$ then takes the form

$$E_{\sigma_t}(t) \overset{\mathcal{D}}{=} \delta_y + \sum_{p \in \pi, i < \nu_p} \eta_{(p)}(p) \sum_{j=1}^{\eta_{(p)}(p)} \delta_{\mathbb{B}^{(p),i}(t-p) - \sqrt{2} \sigma_t}.$$  \hspace{1cm} (9.3.22)

Since, for $1 > \gamma > 1/2$, 

$$\lim_{\sigma \to \infty} \lim_{t \to \infty} \mathbb{P}\left( \forall s \geq \sigma^{-1/2} : Y(t-s) - y + \sqrt{2} \sigma_s s \in \left[-(\sigma_s)^\gamma, (\sigma_s)^\gamma\right] \right) = 1, \hspace{1cm} (9.3.23)$$

if we define the set

$$\mathcal{G}_{\sigma_t} = \left\{ Y : \forall t \geq s \geq \sigma^{-1/2} : Y(t-s) - y + \sqrt{2} \sigma_s s \in \left[-(\sigma_s)^\gamma, (\sigma_s)^\gamma\right] \right\}, \hspace{1cm} (9.3.24)$$

it will suffice to show that, for all $R \in \mathbb{R}_+$,

$$\lim_{\sigma \to \infty} \lim_{t \to \infty} \mathbb{P}\left( \exists p \in \pi, i < \nu_p, j : \mathbb{B}^{(p),i}(t-p) \geq y - R \wedge Y \in \mathcal{G}_{\sigma_t} \right) = 0. \hspace{1cm} (9.3.25)$$

The probability in (9.3.25) is bounded by
\[ \mathbb{P} \left( \exists p \in \pi, t \leq v_p, j : B_{j}^{Y,(p),(j)}(t - p) \geq y - R \wedge Y \in \Theta'_{\sigma} \right) \quad (9.3.26) \]

\[ \leq \mathbb{E} \left[ \int_{0}^{t} \sum_{i=1}^{v_{p}} \mathbb{1}_{\mathbb{Q}_{j}^{Y,(p),(j)}(t - p) > y - R} \mathbb{1}_{Y \in \Theta_{\sigma}} \pi(dp) \right] \]

\[ \leq \mathbb{E} \left[ \int_{0}^{t} \mathbb{E} \left[ \sum_{i=1}^{v_{p}} \mathbb{1}_{\max\mathbb{Q}_{j}^{Y,(p),(j)}(t - p) > y - R} \mathbb{1}_{Y \in \Theta_{\sigma}} | \mathbb{F}_{\infty}^{\pi} \right] \pi(dp) \right] \]

\[ \leq \int_{0}^{t} 2K \mathbb{P} \left( \max_{j} B_{j}^{Y,(t - s)} \geq y - R \wedge Y \in \Theta_{\sigma} \right) ds. \]

Here we used the independence of the offspring BBM and that the conditional probability given the \( \sigma \)-algebra \( \mathbb{F}_{\infty}^{\pi} \) generated by the Poisson process \( \pi \) appearing in the integral over \( \pi \) depends only on \( p \). For the integral over \( s \) up to \( 1/\sigma_{\varepsilon}^{1/2} \), we just bound the integrand by \( 2K \). For larger values, we use the localisation provided by the condition that \( Y \in \Theta_{\sigma} \), to get that the right hand side of (9.3.26) is not larger than

\[ 2K \int_{0}^{\sigma_{\varepsilon}^{-1/2}} ds + 2K \int_{\sigma_{\varepsilon}^{-1/2}}^{\infty} e^{(1 - \sigma_{\varepsilon}^{2})s + \sqrt{2} \sigma_{\varepsilon}(R + (\sigma_{\varepsilon} s)^{\gamma})} ds. \quad (9.3.27) \]

(9.3.27) is by (9.3.9) bounded from above by

\[ 2K \sigma_{\varepsilon}^{-1/2} + 2K \int_{\sigma_{\varepsilon}^{-1/2}}^{\infty} e^{(1 - \sigma_{\varepsilon}^{2})s + \sqrt{2} \sigma_{\varepsilon} R + (\sigma_{\varepsilon} s)^{\gamma}} ds. \quad (9.3.28) \]

From this it follows that (9.3.28) (which does no longer depend on \( t \)) converges to zero, as \( \sigma_{\varepsilon} \uparrow \infty \), for any \( R \in \mathbb{R} \). Hence we see that

\[ \mathbb{P} \left( \sup_{\sigma_{\varepsilon}} ([-R, \infty)) > 1 | \text{supp} \sigma_{\varepsilon} \cap dy \neq 0 \right) \downarrow 0, \quad (9.3.29) \]

uniformly in \( y \geq 0 \), as \( t \) and then \( \sigma_{\varepsilon} \) tend to infinity. Next,

\[ \int_{0}^{\infty} \mathbb{P} \left( \sup_{\sigma_{\varepsilon}} \sigma_{\varepsilon} \cap dy \neq 0 | \max_{k \leq n(t)} x_{k}(t) > \sqrt{2} \sigma_{\varepsilon} t \right) \]

\[ \leq \int_{0}^{\infty} \mathbb{P} \left( \max_{k \leq n(t)} x_{k}(t) \geq \sqrt{2} \sigma_{\varepsilon} t + y | \max_{k \leq n(t)} x_{k}(t) > \sqrt{2} \sigma_{\varepsilon} t \right). \quad (9.3.30) \]

But by Proposition 7.5 in [17] the probability in the integrand converges to \( \exp(-\sqrt{2} \sigma_{\varepsilon} y) \), as \( t \uparrow \infty \). It follows from the proof that this convergence is uniform in \( y \), and hence by dominated convergence, the right-hand side of (9.3.30) is finite. Therefore, (9.3.20) holds. As a consequence, \( A_{\sigma_{\varepsilon}} \) converges to \( \delta_{0} \).

It remains to show that the intensity of the Poisson process converges as claimed. Theorems 1 and 2 of [27] relate the constant \( C(\sigma_{\varepsilon}) \) defined by (9.2.52) to the intensity of the shifted BBM conditioned to exceed the level \( \sqrt{2} \sigma_{\varepsilon} t \) as follows:
\[
\frac{1}{\sqrt{4\pi C(\sigma_e)}} = \lim_{s \to \infty} \mathbb{E} \left[ \max_k \bar{x}_k(s) > \sqrt{2} \sigma_e s \right] \]

(9.3.31)

\[
\begin{align*}
&= \lim_{s \to \infty} \mathbb{E} \left[ \sum_k \mathbb{1}_{\bar{x}_k(s) > \sqrt{2} \sigma_e s} \max_k \bar{x}_k(s) > \sqrt{2} \sigma_e s \right] \\
&= \Lambda_{\sigma_e}((-E, 0]),
\end{align*}
\]

where, by Theorem 7.5 in [17], \( E \) is a exponentially distributed random variable with parameter \( \sqrt{2} \sigma_e \), independent of \( \Lambda_{\sigma_e} \). As we have just shown that \( \Lambda_{\sigma_e} \to \delta_0 \), it follows that the right-hand side tends to one, as \( \sigma_e \to \infty \), and hence \( \tilde{C}(\sigma_e) \to \frac{1}{\sqrt{4\pi}} \). Hence the intensity measure of the PPP appearing in \( \sigma b \sigma_e \) converges to the desired intensity measure \( \frac{1}{\sqrt{4\pi}} \nu_{\sigma_e} e^{-\sqrt{2} x} dx \).

This proves Assertion (iii) of Theorem 9.27.

### 9.3.3 Proof of Proposition 9.29

We prove Proposition 9.29 via convergence of Laplace transforms. Define the Laplace transform of \( \{N_{u_1}(t), \ldots, N_{u_k}(t)\} \),

\[
\mathcal{L}_{u_1, \ldots, u_k}(t, c) = \mathbb{E} \left( \exp \left( - \sum_{l=1}^{k} c_l N_{u_l}(t) \right) \right), \quad c = (c_1, \ldots, c_k) \in \mathbb{R}_+^k,
\]

(9.3.32)

and the Laplace transform \( \mathcal{L}_{u_1, \ldots, u_k}(c) \) of \( \{N_{u_1}, \ldots, N_{u_k}\} \). Proposition 9.29 is then a consequence of the next proposition.

**Proposition 9.31.** For any \( k \in \mathbb{N}, u_1, \ldots, u_k \in \mathbb{R} \) and \( c_1, \ldots, c_k \in \mathbb{R}_+ \)

\[
\lim_{t \to \infty} \mathcal{L}_{u_1, \ldots, u_k}(t, c) = \mathcal{L}_{u_1, \ldots, u_k}(c).
\]

(9.3.33)

The proof of Proposition 9.31 requires two main steps. First, we prove the result for the case of two-speed BBM. This was done in our previous paper [17]. In fact, we will need a slight extension of that result where we allow a slight dependence of the speeds on \( t \). This will be given in the next subsection.

The second step is to show that the Laplace transforms in the general case can be well approximated by those of two-speed BBM. This uses the usual Gaussian comparison argument in a slightly subtle way.
9.3.4 Approximating two-speed BBM.

As we will see later, it is enough to compare the process with covariance function $A$ with processes whose covariance function is piecewise linear with single change in slope. We will produce approximate upper and lower bounds by choosing these in such a way that the covariances near zero and near one are below, resp. above that of the original process. In fact, it is not hard to convince oneself that the following holds.

**Lemma 9.32.** There exist families of functions $\tilde{A}_t$ and $A_t$ that are piecewise linear, continuous functions with a single change point for the derivative, such that $\tilde{A}(0) = A(0) = 0$ and $\tilde{A}(1) = A(1) = 1$. Moreover,

$$\lim_{t \to \infty} \tilde{A}_t(0) = \lim_{t \to \infty} A'_t(0), \quad \text{and} \lim_{t \to \infty} \tilde{A}_t(1) = \lim_{t \to \infty} A'_t(1).$$

(9.3.34)

Moreover,

(i) for all $s$ with $\Sigma^2(s) \in [0, t^{1/3}]$ and $\Sigma^2(s) \in [t - t^{1/3}, t]$,

$$\Sigma^2(s) \geq \Sigma^2(s) \geq \Sigma^2(s).$$

(9.3.35)

(ii) If $A(x) = 0$ on some finite interval $[0, \delta]$, then Eq. (9.3.35) only holds for all $s$ with $\Sigma^2(s) \in [t - t^{1/3}, t]$ while, for $s \in [0, (\delta \wedge \beta) t]$, it holds that

$$\Sigma^2(s) = \Sigma^2(s) = \Sigma^2(s) = 0.$$  

(9.3.36)

Let $\{y_i, i \leq n(t)\}$ be the particles of a BBM with speed function $\Sigma^2$ and let $\{\tilde{y}_i, i \leq n(t)\}$ be particles of a BBM with speed function $\tilde{\Sigma}^2$. We want to show that the limiting extremal processes of these processes coincide. Set

$$\mathcal{N}_u(t) \equiv \sum_{i=1}^{n(t)} \mathbb{1}_{\tilde{y}_i(t) - \tilde{m}(t) > u},$$

(9.3.37)

and

$$\mathcal{N}_u(t) \equiv \sum_{i=1}^{n(t)} \mathbb{1}_{y_i(t) - m(t) > u}.$$  

(9.3.38)

**Lemma 9.33.** For all $u_1, \ldots, u_k$ and all $c_1, \ldots, c_k \in \mathbb{R}_+$, the limits

$$\lim_{t \to \infty} \mathbb{E} \left( \exp \left( - \sum_{i=1}^{k} c_k \mathcal{N}_{u_i}(t) \right) \right)$$

exist. If $A'(1) < \infty$, the two limits coincide with $\mathcal{L}_{u_1, \ldots, u_k}(c)$. 

If $A'(1) = \sigma_e^2 = \infty$, then the two limits in (9.3.39) converges to that in (9.3.40), as $\rho \uparrow \infty$.

**Proof.** We first consider the case when $A'(1) < \infty$. To prove Lemma 9.33, we show that the extremal processes

$$E_t = \sum_{i=1}^{n(t)} \delta_{y_i - \tilde{m}(t)}$$

and

$$E_t = \sum_{i=1}^{n(t)} \delta_{y_i - \tilde{m}(t)}$$

both converge to $E_{\sigma_b, \sigma_e}$, that was defined in (9.3.2). Note that this implies first convergence of Laplace functionals with functions $\phi$ with compact support, while the $M_n(t)$ have support that is unbounded from above. Convergence for these, however, carries over due to tightness.

The assertion in the case when $\sigma_e = \infty$ follows directly from Lemma 9.30. □

### 9.3.5 Gaussian comparison

We now come to the heart of the proof, namely the application of Gaussian comparison to control the process with covariance function $A$ in terms of those with the piecewise linear covariances. We distinguish from now on the expectation with respect to the underlying tree structure and the one with respect to the Brownian movement of the particles.

- $E_n$: expectation w.r.t. Galton-Watson process.
- $E_B$: expectation w.r.t. the Gaussian process conditioned on the $\sigma$-algebra $\mathcal{F}_{\text{tree}}$ generated by the Galton Watson process.

The proof of Proposition 9.31 is based on the following Lemma that compares the Laplace transform $L_{u_1, \ldots, u_k}(t, c)$ with the corresponding Laplace transform for the comparison processes.

**Lemma 9.34.** For any $k \in \mathbb{N}$, $u_1, \ldots, u_k \in \mathbb{R}$ and $c_1, \ldots, c_k \in \mathbb{R}_+$ we have

$$L_{u_1, \ldots, u_k}(t, c) \leq \mathbb{E} \left( \exp \left( - \sum_{i=1}^{k} c_i \tilde{A}_{u_i}(t) \right) \right) + o(1) \quad (9.3.42)$$

$$L_{u_1, \ldots, u_k}(t, c) \geq \mathbb{E} \left( \exp \left( - \sum_{i=1}^{k} c_i \tilde{A}_{u_i}(t) \right) \right) + o(1) \quad (9.3.43)$$

**Proof.** The proofs of (9.3.42) and (9.3.43) are very similar. Hence we focus on proving (9.3.42). We will, however, indicate what has to be changed when proving the lower bound as we go along. For simplicity all overlined names depend on $\Sigma^2$. Corresponding quantities where $\Sigma^2$ is replaced by $\Sigma^2$ are underlined. Set
\[
    f(x(t)) \equiv f(x_1(t), \ldots, x_{n(t)}(t)) \equiv \exp \left( -\sum_{i=1}^{n(t)} \sum_{l=1}^{k} c_l \mathbb{I}_{x_i(t) - \tilde{m}(t) > u_l} \right). \tag{9.3.44}
\]

We want to control
\[
    \mathbb{E}_B \left( \exp \left( -\sum_{i=1}^{k} c_i \mathbb{I}_{x_i(0)} \right) \right) - \mathbb{E}_B \left( \exp \left( -\sum_{i=1}^{k} c_i \mathbb{I}_{y_i(0)} \right) \right)
    = \mathbb{E}_B \left( f(x_1(t), \ldots, x_{n(t)}(t)) \right) - \mathbb{E}_B \left( f(y_1(t), \ldots, y_{n(t)}(t)) \right) \tag{9.3.45}
\]

A straightforward application of the interpolation formula from Lemma 3.1 will produce a term that is very similar to what appears in a second moment estimate. But we already know that this will not give a useful estimate, unless we introduce an appropriate truncation that reflects the typical behaviour of ancestral paths of extrema particles. Thus we have to sneak them in in a clever way. Define as usual the interpolating processes
\[
    x^h_i = \sqrt{h} x_i + \sqrt{1-h} y_i. \tag{9.3.46}
\]

**Remark 9.35.** We understand the interpolating process \( \{x^h_i, i \leq n(t)\} \) as a new Gaussian process with the same underlying branching structure and speed function
\[
    \Sigma^2_h(s) = h\Sigma^2(s) + (1-h)\Sigma^2(s). \tag{9.3.47}
\]

Then, (9.3.45) is equal to
\[
    \mathbb{E}_B \left( \int_0^1 \frac{d}{dh} f(x^h(t)) dh \right), \tag{9.3.48}
\]

where
\[
    \frac{d}{dh} f(x^h(t)) = \frac{1}{2} \sum_{i=1}^{n(t)} \frac{\partial}{\partial x_j} f(x^h_1(t), \ldots, x^h_{n(t)}(t)) \left[ \frac{1}{\sqrt{h}} x_i(t) - \frac{1}{\sqrt{1-h}} y_i(t) \right] \tag{9.3.49}
\]

and \( \frac{\partial}{\partial x_j} f(x^h_1(t), \ldots, x^h_{n(t)}(t)) \) is the weak partial derivative\(^1:\)
\[
    \frac{\partial}{\partial x_j} f(x^h_1(t), \ldots, x^h_{n(t)}(t)) = - \left( \sum_{l=1}^{k} c_l \delta(x^h_1(t) - m(t) - u_l) \right) f(x^h_1(t), \ldots, x^h_{n(t)}(t)). \tag{9.3.50}
\]

\(^1\) In this proof it always suffices to consider weak derivatives since the only non-differentiable function is \( f(x) = 1_{x \geq u} \) which can be monotonously approximated by smooth functions, e.g. replace all indicator functions in \( f \) by \( \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-z^2/2\sigma^2} dz \) rewrite (9.3.45) as in (9.3.48) and then take \( \sigma \uparrow \infty \).
We must introduce the condition on the path of \( x^i_t \) into (9.3.49) at this stage. To do so, we insert into the right-hand side of (9.3.49) a one in the form

\[
1 = \mathbb{1}_{x^i_t \in \mathcal{T}_{t, \Sigma_2^b}} + \mathbb{1}_{x^i_t \notin \mathcal{T}_{t, \Sigma_2^b}},
\]

(9.3.51)

with

\[
I ≡ \left[t(\delta^{<}(t) \wedge \delta^{>}(t)), t(1 - \delta^{<}(t))\right],
\]

(9.3.52)

and \( \mathcal{T}_{t, \Sigma_2^b} \) was defined in (9.3.4). Here \( \delta^{<}, > \) is defined in the same way, but with respect to the speed function \( \Sigma^2 \) instead of \( \Sigma^2 \). We call the two resulting summands \( S^{<} \) and \( S^{>}, \) respectively.

The next lemma shows that \( S^{>}, \) does not contribute to the expectation in (9.3.49), as \( t \to \infty \).

**Lemma 9.36.** With the notation above, we have

\[
\lim_{t \to \infty} \mathbb{E}_n \left( \int_0^1 \mathbb{E}_B(|S^{>}|) dh \right) = 0.
\]

(9.3.53)

The proof of this lemma will be postponed.

We continue with the proof of Lemma 9.34. We are left with controlling

\[
\mathbb{E}_B(S^{<}) = \mathbb{E}_B \left( \frac{1}{2} \sum_{j=1}^{n(t)} \frac{\partial}{\partial x_j} f(x^j_t) \mathbb{1}_{v^i_t \in \mathcal{T}_{t, \Sigma_2^b}} \left[ \frac{x_i(t)}{\sqrt{h}} - \frac{\Sigma(t)}{\sqrt{1-h}} \right] \right).
\]

(9.3.54)

By the definition of \( \mathcal{T}_{t, \Sigma_2^b} \),

\[
\mathbb{1}_{x^i_t \in \mathcal{T}_{t, \Sigma_2^b}} = \mathbb{1}_{v^i_t \in \mathcal{G}(\Sigma_2^b(t))} \left[ \left( \Sigma_2^b(t) \wedge (t - \Sigma_2^b(t)) \right) r \right],
\]

(9.3.55)

where \( \mathcal{G}(\cdot) \) is a Brownian bridge from 0 to 0 in time \( t \), that is independent of \( x^i_t(t) \).

We want to apply a Gaussian integration by parts formula to (9.3.54). However, we need to take care of the fact that each summand in (9.3.54) depends on the whole path of \( \xi_t \) through the term in (9.3.55). However, this is not really a problem, due to the fact that the condition (9.3.55) only depends on the properties of the Brownian bridge which is independent of the endpoint.

By the Gaussian integration by parts formula (3.1.9) we have,

\[
\mathbb{E}_B \left( x_i(t) \frac{\partial}{\partial x_j} f(x^j_t) \mathbb{1}_{v^i_t \in \mathcal{T}_{t, \Sigma_2^b}} \right)
\]

\[
= \sum_{j=1}^{n(t)} \mathbb{E}_B(x_i(t)x_j(t)) \mathbb{E}_B \left( \mathbb{1}_{v^i_t \in \mathcal{T}_{t, \Sigma_2^b}} \frac{\partial^2}{\partial x_j \partial x_i} f(x^j_t) \right),
\]

(9.3.56)

(9.57)

The same formula hold of course with \( x_i \) replaced by \( y_i \). Hence
The term \[ \Sigma \] is replaced by \[ \bar{\Sigma} \] in (9.3.58). Hence, we rewrite (9.3.58) as

\[
\frac{\partial^2 f(x^h(t))}{\partial x_i \partial x_j} = \sum_{i,j=1}^{k} c_{ij} \delta(x^h_i - \bar{m}(t) - u_i) \delta(x^h_j - \bar{m}(t) - u_j) f(x^h_i(t), \ldots, x^h_m(t)).
\]

Introducing

\[
1 = \mathbb{1}_{d(x^h(t), x^f(t)) \in I} + \mathbb{1}_{d(x^h(t), x^f(t)) \notin I},
\]

into (9.3.58) we rewrite (9.3.58) as \((TT) + (T2)\), where

\[
(TT) = \sum_{i,j=1}^{n(t)} \left[ \mathbb{E}_B(x_i(t)x_j(t)) - \mathbb{E}_B(y_i(t)y_j(t)) \right] \mathbb{E}_B \left( \mathbb{1}_{d(x^h(t), x^f(t)) \in I} \prod_{i,j \neq 1} \frac{\partial f(x^h(t))}{\partial x_i \partial x_j} \right),
\]

\[
(T2) = \sum_{i,j=1}^{n(t)} \left[ \mathbb{E}_B(x_i(t)x_j(t)) - \mathbb{E}_B(y_i(t)y_j(t)) \right] \mathbb{E}_B \left( \mathbb{1}_{d(x^h(t), x^f(t)) \notin I} \prod_{i,j \neq 1} \frac{\partial f(x^h(t))}{\partial x_i \partial x_j} \right).
\]

The term \((TT)\) is controlled by the following Lemma.

**Lemma 9.37.** With the notation above, there exists a constant \( \tilde{C} < \infty \), such that

\[
\mathbb{E}_n \left( \int_0^t |(TT)| \, dh \right) \leq \tilde{C} \int_0^t \left| e^{-s + \Sigma^2(s) + O(s^2)} - e^{-s + \bar{\Sigma}^2(s) + O(s^2)} \right| \, ds.
\]

Moreover, we have:

**Lemma 9.38.** If \( \Sigma^2 \) satisfies (A1)-(A3), and \( \bar{\Sigma}^2 \) is as defined in Lemma ??, then

\[
\lim_{t \to \infty} \int_0^t \left| e^{-s + \Sigma^2(s) + O(s^2)} - e^{-s + \bar{\Sigma}^2(s) + O(s^2)} \right| \, ds = 0.
\]

The proofs of these lemmata are technical but fairly straightforward and will not be given here. Details can be found in [18].

Up to this point the proof of (9.3.43) works exactly as the proof of (9.3.42) when \( \Sigma^2 \) is replaced by \( \bar{\Sigma}^2 \). For \((T2)\) and \((TT)\) we have:

**Lemma 9.39.** For almost all realisations of the Galton-Watson process, the following statements hold:

\[
\lim_{t \to \infty} (TT) \leq 0,
\]

and
The proof of this lemma is technical and will be skipped.

From Lemma 9.37, Lemma 9.38, and Lemma 9.39 together with (9.3.54), the bound (9.3.42) follows. As pointed out, using Lemma 9.39, the bound (9.3.43) also follows. Thus, Lemma 9.34 is proved. □

We can now conclude the proof of Proposition 9.31.

Proof (Proof of Proposition 9.31). Taking the limit as \( t \uparrow \infty \) in (9.3.42) and (9.3.43) and using Lemma 9.33 gives, in the case \( A'(1) < \infty \),

\[
\limsup_{t \uparrow \infty} \mathcal{L}_{u_1, \ldots, u_k}(t, c) \leq \mathcal{L}_{u_1, \ldots, u_k}(c)
\]

\[
\liminf_{t \uparrow \infty} \mathcal{L}_{u_1, \ldots, u_k}(t, c) \geq \mathcal{L}_{u_1, \ldots, u_k}(c)
\]  

(9.3.67)

Hence \( \lim_{t \uparrow \infty} \mathcal{L}_{u_1, \ldots, u_k}(t, c) \) exists and is equal to \( \mathcal{L}_{u_1, \ldots, u_k}(c) \). In the case \( A'(1) = \infty \), the same result follows if in addition we take \( \rho \uparrow \infty \) after taking \( t \uparrow \infty \). This concludes the proof of Proposition 9.31. □
References

References

References

## Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>almost sure convergence</td>
<td>23</td>
</tr>
<tr>
<td>ancestral path</td>
<td>58</td>
</tr>
<tr>
<td>asymptotics of upper tails</td>
<td>117</td>
</tr>
<tr>
<td>Berman condition</td>
<td>33</td>
</tr>
<tr>
<td>boundary</td>
<td></td>
</tr>
<tr>
<td>of Galton-Watson process</td>
<td>53</td>
</tr>
<tr>
<td>branching Brownian motion</td>
<td>57</td>
</tr>
<tr>
<td>variable speed</td>
<td>151</td>
</tr>
<tr>
<td>branching process</td>
<td>52</td>
</tr>
<tr>
<td>Brownian bridge</td>
<td>58, 72, 92, 111</td>
</tr>
<tr>
<td>basic estimates</td>
<td>92</td>
</tr>
<tr>
<td>complete Poisson convergence</td>
<td>30</td>
</tr>
<tr>
<td>convergence theorem</td>
<td>64</td>
</tr>
<tr>
<td>covariance</td>
<td>47, 57</td>
</tr>
<tr>
<td>Cox process</td>
<td>55</td>
</tr>
<tr>
<td>derivative martingale</td>
<td>66, 67, 122</td>
</tr>
<tr>
<td>Dynkin’s theorem</td>
<td>17</td>
</tr>
<tr>
<td>embedding in $\mathbb{R}$</td>
<td></td>
</tr>
<tr>
<td>of Galton-Watson tree</td>
<td>137</td>
</tr>
<tr>
<td>entropic repulsion</td>
<td>59, 75, 95</td>
</tr>
<tr>
<td>ergodic theorem</td>
<td>70</td>
</tr>
<tr>
<td>exceedances</td>
<td></td>
</tr>
<tr>
<td>point process</td>
<td>28</td>
</tr>
<tr>
<td>existence</td>
<td>120</td>
</tr>
<tr>
<td>existence of solutions</td>
<td>75</td>
</tr>
<tr>
<td>extremal distributions</td>
<td>2</td>
</tr>
<tr>
<td>extremal process</td>
<td>15</td>
</tr>
<tr>
<td>extremal types</td>
<td>10</td>
</tr>
<tr>
<td>extremal types theorem</td>
<td>10</td>
</tr>
<tr>
<td>extreme value theory</td>
<td>45</td>
</tr>
<tr>
<td>F-KPP equation</td>
<td>61</td>
</tr>
<tr>
<td>F-KPP equation linear</td>
<td>89</td>
</tr>
<tr>
<td>Feynman-Kac formula</td>
<td>83, 99</td>
</tr>
<tr>
<td>Feynman-Kac representation</td>
<td>71</td>
</tr>
<tr>
<td>finite mass condition</td>
<td>104</td>
</tr>
<tr>
<td>first moment estimate</td>
<td>59</td>
</tr>
<tr>
<td>FKG inequality</td>
<td>110</td>
</tr>
<tr>
<td>Fréchet distribution</td>
<td>10</td>
</tr>
<tr>
<td>Galton-Watson process</td>
<td>52, 137</td>
</tr>
<tr>
<td>Galton-Watson tree</td>
<td>52</td>
</tr>
<tr>
<td>embedding in $\mathbb{R}$</td>
<td>137</td>
</tr>
<tr>
<td>Gaussian comparison</td>
<td>34, 184</td>
</tr>
<tr>
<td>Gaussian integration by parts</td>
<td>35</td>
</tr>
<tr>
<td>Gaussian process</td>
<td>47, 57</td>
</tr>
<tr>
<td>Gaussian sequence</td>
<td></td>
</tr>
<tr>
<td>stationary</td>
<td>33</td>
</tr>
<tr>
<td>Gaussian tail bounds</td>
<td>3</td>
</tr>
<tr>
<td>generalised random energy model</td>
<td>44, 47</td>
</tr>
<tr>
<td>Gibbs measure</td>
<td>45</td>
</tr>
<tr>
<td>Girsanov formula</td>
<td>96</td>
</tr>
<tr>
<td>GREM</td>
<td>44, 47, 151</td>
</tr>
<tr>
<td>Gumbel distribution</td>
<td>4, 47, 65</td>
</tr>
<tr>
<td>random shift</td>
<td>55</td>
</tr>
<tr>
<td>Hamiltonian</td>
<td>43</td>
</tr>
<tr>
<td>Hamming distance</td>
<td>43</td>
</tr>
<tr>
<td>integration by parts</td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>35</td>
</tr>
<tr>
<td>intensity measure</td>
<td>18</td>
</tr>
<tr>
<td>interpolating process</td>
<td>184</td>
</tr>
<tr>
<td>interpolation</td>
<td>34</td>
</tr>
<tr>
<td>inverse</td>
<td></td>
</tr>
<tr>
<td>left continuous</td>
<td>5</td>
</tr>
<tr>
<td>Kahane's theorem</td>
<td>36</td>
</tr>
<tr>
<td>Index</td>
<td>Index</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Kallenberg’s theorem, 25</td>
<td>of BBM, 120</td>
</tr>
<tr>
<td>Khintchine’s theorem, 7</td>
<td>Poisson, 19</td>
</tr>
<tr>
<td>Kolmogorov’s theorem, 79</td>
<td>Poisson point process, 19</td>
</tr>
<tr>
<td>labelling, 53</td>
<td>random energy model, 45</td>
</tr>
<tr>
<td>Laplace functional, 18. 50 63</td>
<td>relatively compact, 22</td>
</tr>
<tr>
<td>convergence, 120</td>
<td>REM, 45</td>
</tr>
<tr>
<td>Laplace transform, 18</td>
<td>second moment method, 46 47</td>
</tr>
<tr>
<td>law of maximum, 69</td>
<td>truncated, 49</td>
</tr>
<tr>
<td>REM, 46</td>
<td>Skorokhod’s theorem, 23</td>
</tr>
<tr>
<td>lexicographic distance, 44</td>
<td>Slepian’s lemma, 36</td>
</tr>
<tr>
<td>linear F-KPP equation, 71 75 89</td>
<td>spin glass, 43</td>
</tr>
<tr>
<td>localisation of paths, 59</td>
<td>standard conditions, 63 71</td>
</tr>
<tr>
<td>max-stability, 5</td>
<td>stationary Gaussian sequence, 33</td>
</tr>
<tr>
<td>max-stable distribution, 8</td>
<td>stretched</td>
</tr>
<tr>
<td>maxima of iid sequences, 1</td>
<td>more, 80</td>
</tr>
<tr>
<td>maximum principle, 76</td>
<td>supercritical, 54</td>
</tr>
<tr>
<td>mild solution, 62</td>
<td>tightness, 24 120 177</td>
</tr>
<tr>
<td>more stretched, 80</td>
<td>travelling wave, 63</td>
</tr>
<tr>
<td>multi-index, 53</td>
<td>ultrametric, 44 53</td>
</tr>
<tr>
<td>labelling, 53</td>
<td>uniqueness of solutions, 75</td>
</tr>
<tr>
<td>normal comparison, 34</td>
<td>vague convergence, 21</td>
</tr>
<tr>
<td>normal sequence, 33</td>
<td>variable speed, 151</td>
</tr>
<tr>
<td>order statistics, 11</td>
<td>weak convergence, 5</td>
</tr>
<tr>
<td>particle number, 60</td>
<td>of point processes, 20</td>
</tr>
<tr>
<td>partition function, 68</td>
<td>Weibull distribution, 10</td>
</tr>
<tr>
<td>Picard iteration, 75</td>
<td>Yule process, 52</td>
</tr>
<tr>
<td>point measure, 15</td>
<td></td>
</tr>
<tr>
<td>point process, 15 16 63</td>
<td></td>
</tr>
</tbody>
</table>